Improved Spatial Modulation Techniques for Wireless Communications

A Thesis Submitted
to the College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Electrical and Computer Engineering
University of Saskatchewan

by

Binh T. Vo

Saskatoon, Saskatchewan, Canada

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, it is agreed that the Libraries of this University may make it freely available for inspection. Permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professors who supervised this thesis work or, in their absence, by the Head of the Department of Electrical and Computer Engineering or the Dean of the College of Graduate and Postdoctoral Studies at the University of Saskatchewan. Any copying, publication, or use of this thesis, or parts thereof, for financial gain without the written permission of the author is strictly prohibited. Proper recognition shall be given to the author and to the University of Saskatchewan in any scholarly use which may be made of any material in this thesis.

Request for permission to copy or to make any other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Electrical and Computer Engineering
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A9
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

Abstract

Transmission and reception methods with multiple antennas have been demonstrated to be very useful in providing high data rates and improving reliability in wireless communications. In particular, spatial modulation (SM) has recently emerged as an attractive transmission method for multiple-antennas systems due to its better energy efficiency and lower system complexity. This thesis is concerned with developing transmission techniques to improve the spectral efficiency of SM where antenna/subcarrier index involves in conveying information bits.

In the first part of the thesis, new transmission techniques are developed for SM over frequency-flat fading channels. The first proposed scheme is based on a high-rate space-time block code instead of using the classical Alamouti STBC, which helps to increase the spectral efficiency and achieve a transmit diversity order of two. A simplified maximum likelihood detection is also developed for this proposed scheme. Analysis of coding gains and simulation results demonstrate that the proposed scheme outperforms previously-proposed SM schemes at high data transmission rates. Then, a new space-shift keying (SSK) modulation scheme is proposed which requires a smaller number of transmit antennas than that required in the bi-space shift keying (BiSSK). Such a proposed SSK-based scheme is obtained by multiplexing two in-phase and quadrature generalized SSK streams and optimizing the carrier signals transmitted by the activated antennas. Performance of the proposed scheme is compared with other SSK-based schemes via minimum Euclidean distance analysis and computer simulation. The third scheme proposed in this part is an improved version of quadrature SM (QSM). The main feature of this proposed scheme is to send a second constellation symbol over the in-phase and quadrature antenna dimensions. A significant performance advantage of the proposed scheme is realized at the cost of a slight increase in the number of radio-frequency (RF) chains. Performance comparisons with the most recent SM schemes confirm the advantage of the proposed scheme. The last contribution of the first part is an optimal constellation design for QSM to minimize the average probability of error. It is shown that, the error performance of QSM not only depends on the Euclidean distances between the amplitude phase modulation (APM) symbols and the energies of APM symbols, but also on the

in-phase and quadrature components of the QSM symbols. The analysis of the union bound of the average error probability reveals that at a very large number of transmit antennas, the optimal constellations for QSM converge to a quadrature phase shift keying (QPSK) constellation. Simulation results demonstrate the performance superiority of the obtained constellations over other modulation schemes.

In the second part of the thesis, the applications of SM in frequency-selective fading channels are studied. First, a new transmission scheme that employs SM for each group of subcarriers in orthogonal frequency-division multiplexing (OFDM) transmission is investigated. Specifically, OFDM symbols in each group are passed through a precoder to maximize the diversity and coding gains, while SM is applied in each group to convey more information bits by antenna indices. Performance analysis and simulation results are carried out to demonstrate the superiority of the proposed scheme over a previously-proposed combination of SM and OFDM. Next, the performance of OFDM based on index modulation and a flexible version of OFDM, knows as OFDM with multiple constellations, is compared for both case of "no precoding" and "with precoding" of data symbols. It is shown that the precoded OFDM with multiple constellations outperforms precoded-IM based OFDM systems over frequency-selective fading channels.

The last part of the thesis investigates a multiuser downlink transmission system based on in-phase and quadrature space-shift keying modulation and precoding to reduce the minimum number of transmit antennas while keeping the complexity of the receiver low. In addition to the maximum likelihood (ML) detection, the low complexity zero forcing (ZF) receiver is also studied. Theoretical upper bounds for the error probabilities of both ML and ZF receivers are obtained and corroborated with simulation results.

Acknowledgments

It is a pleasure to take this opportunity to convey thanks to those who made this thesis possible with their kindness and support.

First, I would like to express my very great appreciation to my supervisor, Professor Ha H. Nguyen, for his patient guidance, enthusiastic encouragement and useful critiques during the planning and development of my research work at the University of Saskatchewan. It has truly been my honour and rewarding experience to work under his supervision. Without him, this thesis would not have been completed.

I would also like to express my gratitude to Professors Joseph E. Salt, Rajesh Karki, Oon-Doo Baik, Seok-Bum Ko from the University of Saskatchewan, and Professor Geoffrey Messier from the University of Calgary, for serving in my doctoral committee and for reviewing and evaluating this thesis. Their insightful comments and suggestions have improved the quality of this thesis.

My deepest love and gratitude go to my family for supporting my studies with enthusiasm and encouragement. My special thanks are extended to all my past and current members of our research group for sharing their knowledge and invaluable assistance.

Table of Contents

List of Abbreviations

ABEP	Average bit error probability
APM	Amplitude phase modulation
AWGN	Additive white Gaussian noise
BER	Bit error rate
BiGSSK	Bi-generalized space-shift keying
BiSSK	Bi-space shift keying
BPSK	Binary phase-shift keying
CDMA	Code-division multiple access
CP	Cyclic prefix
D-BLAST	Diagonal Bell Labs layered space time
DFT	Discrete Fourier transform
ESM	Enhanced spatial modulation
FFT	Fast Fourier transform
GLCP	Grouped linear constellation precoding
GSM	Generalized spatial modulation
GSSK	Generalized space-shift keying
HR-STBC-SM	High-rate space-time block-coded spatial modulation
IAS	Inter-antenna synchronization
ICI	Inter-channel interference

IDFT	Inverse discrete Fourier transform
IFFT	Inverse fast Fourier transform
IM	Index modulation
IQSM	Improved quadrature spatial modulation
ISI	Inter-symbol interference
LCP	Linear constellation precoding
LOS	Light-of-sight
LSSK	Layered space shift keying
MAC	Media access control
MIMO	Multiple-input multiple-output
MISO	Multiple-input single-output
ML	Maximum likelihood
MMSE	Minimum mean square error
MRC	Maximum-ratio combining
MU-PQSSK	Multiuser precoding aided quadrature space shift keying
OFDM	Orthogonal frequency-division multiplexing
OSTBC	Orthogonal space-time block code
P/S	Parallel-to-serial
pcu	Per channel use
PEP	Pairwise error probability
PSK	Phase-shift keying

QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
QSM	Quadrature spatial modulation
RF	Radio frequency
Rx	Receiver
SC	Selection combining
SER	Symbol error rate
SIMO	Single-input multiple-output
SISO	Single-input single-output
SM	Spatial modulation
SMUX	Spatial multiplexing
SNR	Signal-to-noise ratio
SSK	Space-shift keying
STBC-SM	Space-time block-coded spatial modulation
STC	Space-time coding
Tx	Transmitter
V-BLAST	Vertical Bell Labs layered space time
ZF	Zero forcing

List of Figures

List of Tables

1. Introduction

1.1 Motivation

The explosive growth in the use of mobile devices makes wireless communications an important part of human's daily activities. The next and future generations of wireless communication systems require higher transmission rates which implies larger bandwidths. However, the bandwidth of a radio-frequency channel is a finite and expensive resource. To meet the demand of higher transmission rates, better resource management and improved transmission technologies are needed. Multiple-input multiple-output (MIMO) technology that exploits the use of multiple antennas at the transmitter and/or receiver has been demonstrated to be very useful for providing high-data rate transmission [4–6].

Before explaining the benefits of using multiple antennas over a single antenna, it is instructive to understand the main challenges in dealing with a wireless communication channel, which might be considered as the only element that the designer does not have a full control over. For wireless communications, *fading* refers to the variation that the signal level experiences due to the propagation characteristics of a wireless channel. Experimental measurements show that the instantaneous received power over time resembles realizations of a random process. As power fluctuates randomly, it may approach a vanishingly small level, in which case the wireless channel is said to be in a deep fade. The random nature of the wireless fading channel is a major difference when compared to a wired channel, which is generally modeled as deterministic. In general, the fading effects caused by wireless propagation can be categorized in three categories: (i) large-scale path loss effect, modeled through a signal envelope that decays with distance squared; (ii) large-to-medium scale slowly-varying shadowing effect, modeled by a random channel amplitude that follows a log-

normal distribution; and (iii) small-scale fast-varying effect, modeled as a random channel amplitude adhering to a Rice distribution if the light-of-sight (LOS) component is present, or to a Rayleigh distribution if the LOS component is absent, or more generally, to a Nakagami distribution which can approximate a Rayleigh or Rice distribution [4, 5].

To reduce the effect of small-scale fading, a technique called *diversity* is widely used in modern wireless communication systems. Roughly speaking, diversity is implemented by sending copies of signal over independent resource dimensions to reduce the probability of occurrence of a deep fade. The performance gain achieved by a diversity technique is usually quantified by the *diversity gain*, which is basically the decaying order of the error probability curve versus the signal-to-noise ratio (SNR). Depending on the system resource that can be used to carry the multiple signal copies, diversity can be categorized into the following types [4, 5, 7]:

- *Time diversity*: Time diversity is implemented by sending copies of a signal over multiple time slots over which the channel responses are independent. Specifically, time diversity can be implemented by simply repeating the same signal over different time slots, a method called *repetition code*. More efficient schemes include using rotation matrix or interleaved channel code which can guarantee that different errors are related to multiple independent channels.

- *Frequency diversity*: Frequency diversity can be obtained in a system that operates over a wide-band channel (i.e., with a large enough bandwidth) so that different propagation paths can be distinguished. Specifically, frequency diversity can simply be implemented by sending the same information symbol on multiple carriers that are well separated in the frequency domain. In that way, each symbol replica on a different carrier will experience different fading effect. A well-studied approach to achieve frequency diversity is using orthogonal frequency-division multiplexing (OFDM) combined with channel coding or precoding. Another scheme that exploits frequency diversity is the Rake receiver used in code-division multiple access (CDMA) systems [8].

- *Space diversity*: Space diversity is implemented by installing multiple antennas at the transmitter and/or receiver. If multiple antennas are placed at the transmitter, a

method called space-time coding (STC) is usually used to achieve diversity gain and the method is called *transmit diversity*. On the other hand, if multiple antennas are installed at the receiver, the diversity gain can be achieved by implementing combining schemes such as the maximum-ratio combining (MRC), equal gain combining or selection combining. These methods are called *receive diversity*. Of course, multiple antennas can also be implemented at both the transmitter and receiver, leading to the popular multiple-input multiple-output (MIMO) system.

As discussed above, time diversity is obtained by expanding the time window, while frequency diversity requires a large transmission bandwidth. By deploying multiple antennas, space diversity does not require time or bandwidth expansion. This is the key advantage of a MIMO system as compared to a single-antenna system. The diversity gain achieved by a MIMO system improves detection quality. However, the spectral efficiency, measured as data rate per unit bandwidth, of a conventional MIMO system can be low. For instant, in orthogonal space-time coding, the spectral efficiency is one symbol per channel use when the number of transmit antennas, N_t, equals two (which is achieved by using the famous Alamouti space-time code) and smaller than one when $N_t > 2$. In contrast, the spectral efficiency is always one symbol per channel use in a single-input single-output (SISO) system. To overcome the rate limitation of space-time coding, another MIMO approach was investigated by making use of the large number of antennas to maximize the spectral efficiency. An well-known example of this MIMO approach is V-BLAST (Vertical Bell Labs Layered Space Time) [9]. This scheme demultiplexes a data stream into a number of substreams to be transmitted on multiple transmit antennas. However, a challenge of implementing this scheme is to handle inter-channel interference (ICI) and maintain inter-antenna synchronization (IAS) between data streams from different transmit antennas. This scheme also requires a high complexity of decoding at the receiver. More importantly, the transmit-diversity gain is not achieved with V-BLAST.

The advantage of a MIMO system over a traditional single-antenna system is that it can multiply the capacity of a wireless connection without requiring more bandwidth. The more antennas the transmitter/receiver is equipped with, the larger the number of possible

signal paths and the better the performance in terms of data rate and link reliability become. Thus, an obvious approach regarding the enhancement of MIMO capacity is deploying large-scale antenna systems, which are more commonly known as "massive MIMO". Conventional MIMO systems typically use two or four antennas, while massive MIMO systems can be equipped with hundreds of antennas. If properly designed, massive MIMO brings huge improvements in throughput and energy efficiency. Other benefits of massive MIMO include the extensive use of inexpensive low-power components, reduced latency, simplification of the media access control (MAC) layer, and robustness to interference and intentional jamming [10].

While MIMO technology has become an essential part of modern wireless communication systems, there are scenarios that wireless devices cannot support multiple antennas due to size, cost, and/or hardware limitations. Furthermore, as the number of antennas increases, the actual MIMO performance falls far behind the theoretical gains. Cooperative MIMO uses distributed antennas on different radio devices to achieve close to the theoretical gains of MIMO. The basic idea of cooperative MIMO is to group multiple devices into a virtual antenna array to achieve MIMO communications. A cooperative MIMO transmission involves multiple point-to-point radio links, including links within a virtual array and possibly links between different virtual arrays. The disadvantages of cooperative MIMO come from the increased system complexity and the large signaling overhead required for supporting device cooperation. The advantages of cooperative MIMO, on the other hand, are its capability to improve the capacity, cell edge throughput, coverage, and group mobility of a wireless network at a low cost. In recent years, cooperative MIMO technologies have been adopted into the mainstream of wireless communication standards [11].

Spatial modulation (SM) technology has been proposed as an improved and flexible version of the conventional MIMO communications [12, 13]. Specifically, while the conventional MIMO communication system transmits multiple data streams on all available antennas, spatial modulation in MIMO (SM-MIMO) only transmits on a subset of available antennas. By doing so, ICI and IAS are removed or much reduced since the number of data streams is less. A unique feature of SM is that additional information bits can be sent by properly

indexing (i.e., switching) among antenna subsets. With this feature, the spectral efficiency of SM-MIMO will be higher than one symbol per channel use. SM-MIMO can still realize the spatial multiplexing gain and transmit diversity gain of the conventional MIMO communications, while requiring only a few (possibly a single) activated antenna elements (single-RF front-end) at the transmitter at any modulation instant [14]. By using a limited number of RF chains, the signal processing and circuitry complexity are much reduced, leading to an improved energy efficiency of the whole communication system [15]. In addition to the spectrum efficiency, energy efficiency has been recognized as an equally-important performance indicator to guide the design and optimization of transmission technologies and protocols for next generation wireless networks. In essence, by explicitly taking into account the energy consumption and the system's complexity, energy efficiency provides an indication of the throughput per unit energy [16]. Recent analytical and simulation studies have shown that SM-MIMO techniques have the inherent potential of outperforming many state-of-the-art MIMO schemes, provided that a sufficiently-large number of antenna elements is available at the transmitter, while just few of them are simultaneously active [17].

1.2 Research Objectives

The main objective of this research is to develop improved transmission techniques for wireless communication systems [1]. The new transmission techniques are inspired by recent research activities in the general area of spatial modulation in MIMO communications.

From the overview of research activities dedicated to the analysis and design of SM-MIMO communication, it is noted that, most transmission schemes of SM-MIMO under flat fading channels have a low spectral efficiency. Motivated by this observation, the first problem to be considered in this thesis is to develop transmission techniques based on the concept of spatial modulation under flat fading channels to enhance the transmission rate at a negligible performance loss as well as allowing simple transmitter and receiver implementation.

Orthogonal frequency-division multiplexing (OFDM) is a multi-carrier transmission method which is commonly used for frequency-selective fading channels due to its ability to transform a wide-band channel into a set of multiple narrow-band channels. Motivated by the

potential of SM and OFDM and the fact that only limited research considered the combination of these two techniques, the second part of this research focuses on the joint design of SM and OFDM. The majority of research on SM schemes was performed assuming flat fading channels. If those schemes are directly applied to SM-OFDM systems, they may not be efficient or even valid. Furthermore, direct application of SM for OFDM systems does not realize the inherent frequency diversity gain since information symbols are not spread over independently-faded subcarriers. Although frequency diversity combining with OFDM has been well studied in point-to-point communications, this part of the research investigates how to exploit frequency diversity gain in SM-OFDM systems.

Many SM schemes have been studied by implicitly assuming a single-user transmission. However, this network scenario is quite restrictive for a typical cellular deployment, where many users simultaneously transmit over the same frequency bands and/or time slots. As such, the final part of this research develops transmission techniques that are based on spatial modulation and robust to multiple-access interference.

1.3 Organization of the Thesis

This thesis includes six chapters. The first chapter gives the motivation of the research.

Chapter 2 contains the background on point-to-point wireless communications. This chapter first describes single-antenna communication systems, the block diagram of transmitter and receiver, the wireless channel model and modulation and demodulation techniques. Next, it discusses multiple-antenna communications systems, receive diversity with combining methods, transmit diversity with space-time coding, and spatial multiplexing techniques. Spatial modulation is then introduced under a frequency-flat fading channel. The transmitter and receiver structures as well as the advantages and disadvantages of spatial modulation are discussed in detail. Finally, the chapter presents different methods in applying spatial modulation for frequency-selective fading channels.

Chapter 3 presents various novel transmission schemes based on spatial modulation for frequency-flat fading channels. First, a high-rate space-time block-coded spatial modulation

with two active antennas is introduced in which the data rate is increased by making use of the high-rate space time block code design. The simplified maximum-likelihood (ML) detection is also developed to allow for low complexity detection. Next, a scheme which doubles the transmission rate of the generalized space-shift keying (SSK)[1] while still enjoying low complexity is introduced by employing both in-phase and quadrature components. The chapter then investigates quadrature spatial modulation (QSM) and presents an improved QSM scheme that achieves a higher data rate. Finally, constellation design for QSM is considered to optimize the performance.

In Chapter 4, the application of spatial modulation in frequency-selective fading channels is studied. In particular, a novel scheme, called spatial modulation OFDM with grouped linear constellation precoding (SM-OFDM-GLCP), is presented where subcarriers are divided into groups and information symbols are spread across subcarriers in each group to maximize the diversity and coding gains. The spatial modulation is applied to select an antenna to transmit OFDM symbols of the group. Performance analysis and numerical results are presented to determine the diversity and coding gains of the proposed scheme. The last part of the chapter makes a comparison between index-modulation (IM) based OFDM and precoded-OFDM with multiple constellations to give a clear understanding of the merits of the IM-based OFDM systems.

Chapter 5 investigates spatial modulation under multiuser scenario. The special form of spatial modulation, SSK, is applied for downlink communication from a base station to multiple users. A precoding scheme is applied at the base station to mitigate the multiuser interference. Then, a novel transmission scheme taking the advantage of combining inphase and quadrature SSK is presented to allow a smaller number of transmit antennas at the base station. The maximum likelihood and zero forcing detection methods are also developed for the proposed scheme.

Chapter 6 concludes this thesis by summarizing the contributions and suggesting potential research problems for further studies.

[1]SSK is a special form of spatial modulation in which only the antenna index is used to convey the information.

2. Background

This chapter provides necessary background which should be helpful in understanding the main research contributions presented in subsequent chapters of the thesis. The first part of this chapter explains a point-to-point wireless communication system in which the transmitter and the receiver are equipped with a single antenna. Here, the block diagram of the transmitter and receiver, the channel model, and modulation and demodulation techniques are described in detail. It then introduces multiple-antennas systems (with multiple antennas equipped at the transmitter and/or receiver) that achieve the diversity gain by using different signal combining methods at the receiver and space-time codes at the transmitter. Next, the conventional SM system is presented for a frequency-flat fading channel and its advantages and disadvantages are discussed. Finally, the chapter explains orthogonal frequency-division multiplexing (OFDM) and how it can be combined with SM for operating over a frequency-selective fading channel.

2.1 Single-Antenna Wireless Communication Systems

Figure 2.1 depicts the block diagram of a wireless communication system using a single antenna at both the transmitter (Tx) and the receiver (Rx). At the transmitter, the information bits are converted to symbols based on some modulation scheme. In wireless communications, the modulation involves changing the amplitude and/or phase of a sinusoidal carrier with frequency f_c. Such modulation schemes are often represented by two-dimensional constellation plots such as those in Figure 2.2 on page 10. In general, each constellation plot has M points and is called an M-ary constellation. Usually $M = 2^\lambda$ for some integer λ, which also means that each constellation point can carry λ bits. The two

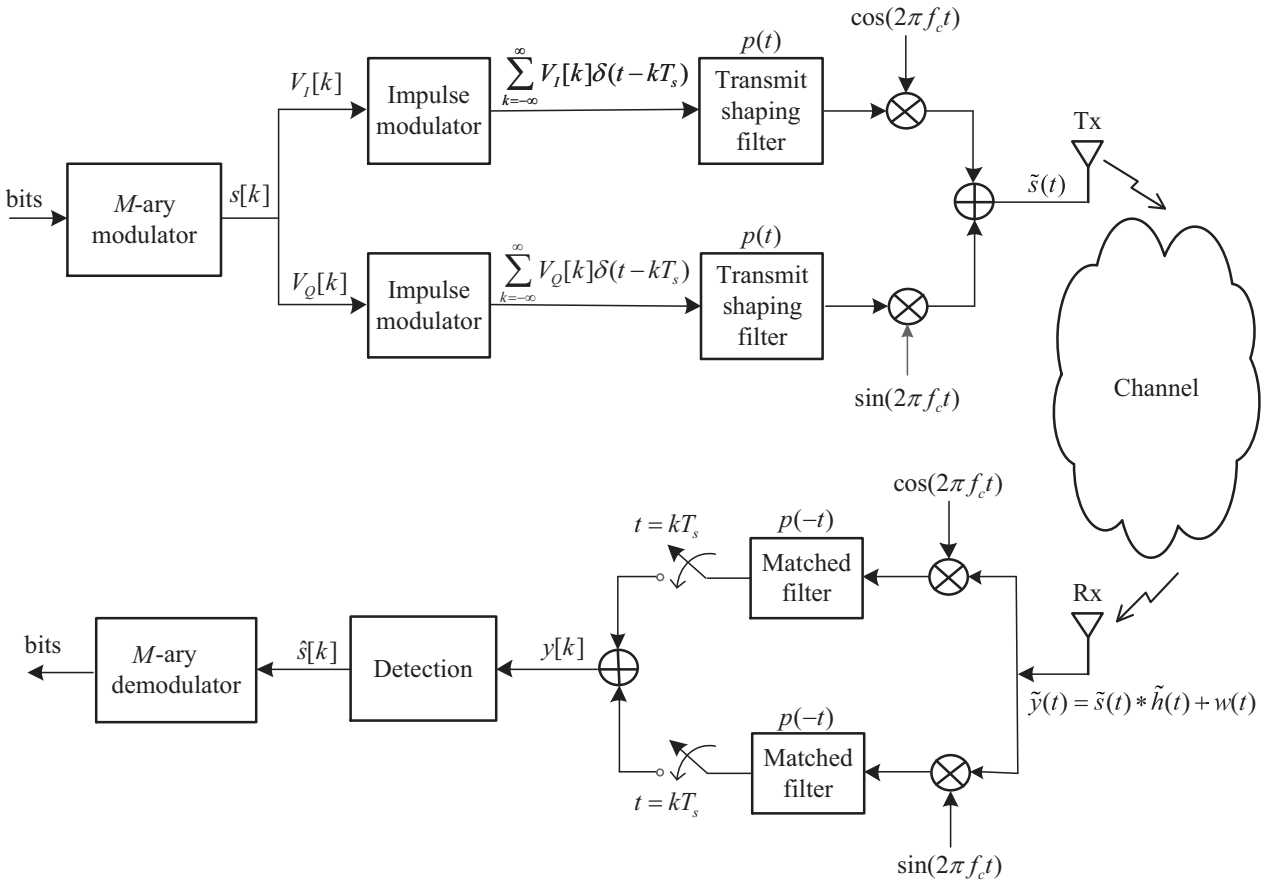

Figure 2.1 Block diagram of a single-antenna wireless communication system.

examples in Figure 2.2 are for the cases of $M = 4$ and $M = 16$ and the mapping from $\lambda = 2$ bits and $\lambda = 4$ bits to the constellation points are also indicated. These two constellations are known as quadrature amplitude modulation (QAM) since the transmitted signal is generated by modulating (i.e., varying) the amplitudes of the two quadrature carriers: $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$. Specifically, a constellation point in the QAM constellation can be represented as a complex number as $s[k] = V_I[k] + jV_Q[k]$, where k is the symbol index. Such a complex symbol corresponds to the following radio-frequency (RF) signal:

$$\tilde{s}(t) = \left[\sum_{k=-\infty}^{\infty} V_I[k]p(t - kT_s)\right]\cos(2\pi f_c t) + \left[\sum_{k=-\infty}^{\infty} V_Q[k]p(t - kT_s)\right]\sin(2\pi f_c t) \qquad (2.1)$$

where $p(t)$ is the impulse response of the pulse shaping filter and T_s is the symbol period. It is pointed out that the 4-QAM constellation is more commonly referred to as quadrature phase-shift keying (QPSK). This is because the 4-QAM signal can be represented as a constant-amplitude sinusoid with four different phases, namely $\cos(2\pi f_c t + n(\pi/4))$, $n = 1, 2, 3, 4$.

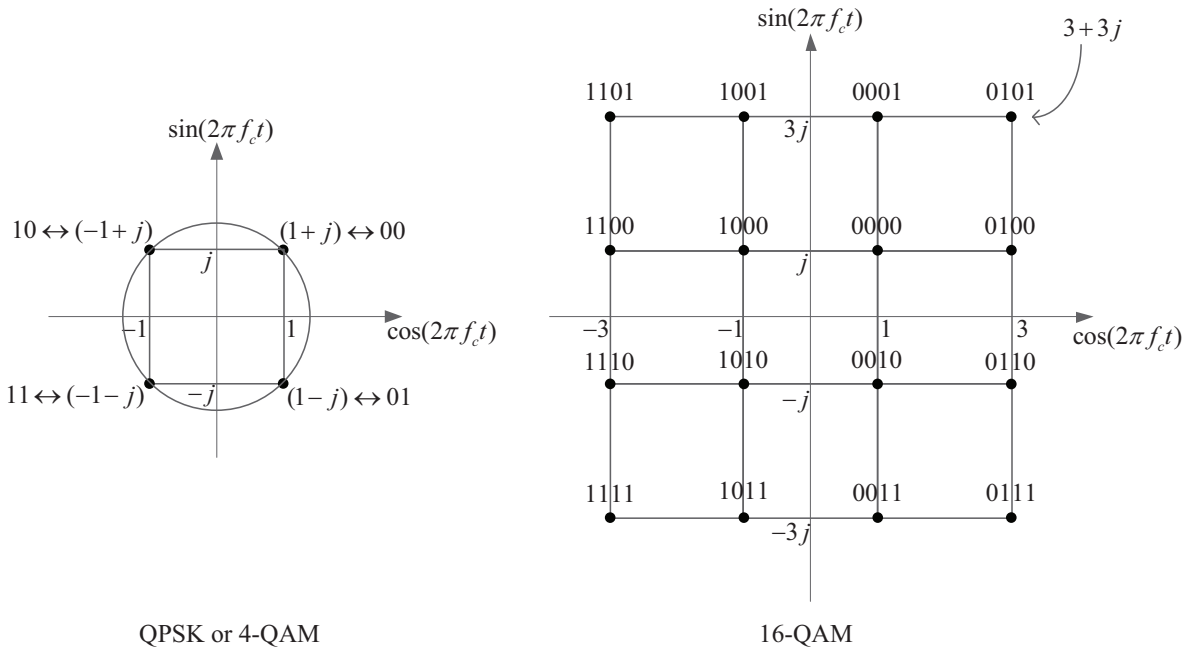

Figure 2.2 Illustration of QPSK ($M = 4$) and 16-QAM ($M = 16$) constellations.

Next, the RF signal is propagated through the wireless channel and would typically be affected by the fading phenomena. Hence, it is important to look at the mathematical model of a fading channel. For flat-fading channels, which are applicable for narrowband systems, the channel impulse response $\tilde{h}(t)$ is represented by a gain, or with a single filter tap (or coefficient). In addition, since most of the processing is actually done at the baseband, the baseband representation of the channel coefficient is used. This coefficient is modeled as a complex random variable whose distribution depends on the nature of the radio propagation environment. Typical distributions are Rayleigh, Rician, Nakagami [4,5]. Here, the Rayleigh flat-fading model is used since it is widely viewed as the worst-case wireless channel model. Let $h[k]$ represent the channel coefficient. Then, the Rayleigh fading $h[k]$ is modeled as a complex Gaussian random variable with zero mean and variance σ_h^2. In fact, the name "Rayleigh" fading comes from the distribution of the envelope $\eta = |h[k]|$, which is a Rayleigh distribution:

$$f_\eta(\eta) = \frac{2\eta}{\sigma_h^2} \exp\left(\frac{-\eta}{\sigma_h^2}\right), \quad \eta \geq 0 \tag{2.2}$$

The received signal is processed at the receiver whose structure can be modeled as shown in Figure 2.1. The signal from the receive antenna is further disturbed by additive white

Gaussian noise $w(t)$. Then, the received RF signal $\tilde{y}(t) = \tilde{s}(t) * \tilde{h}(t) + w(t)$ is down-converted and passed through the match filter to obtain the continuous-time baseband signal. The signal is then sampled at the symbol rate, i.e., at $t = kT_s$, to obtain the discrete-time baseband signal as

$$y[k] = h[k]s[k] + w[k] \tag{2.3}$$

where $h[k] \sim \mathcal{CN}(0, \sigma_h^2)$ is the channel fading coefficient and $w[k] \sim \mathcal{CN}(0, N_0)$ is the white Gaussian noise component at the receiver. To arrive at (2.3), it is assumed that the transmitter and receiver are synchronized and that the transmit shaping filter and the matched filter are designed to satisfy the Nyquist criterion for zero inter-symbol interference (ISI).

To understand the performance of the communication system in (2.3) under the influence of fading, let's first review the error probability of the following system over an additive white Gaussian noise (AWGN) channel without fading:

$$y[k] = s[k] + w[k] \tag{2.4}$$

For antipodal signaling with binary phase-shift keying (BPSK), the symbol to be transmitted is either a or $-a$. The maximum likelihood (ML) detection of the system in (2.4) when a single symbol is transmitted is

$$\hat{s}[k] = \arg \min_{s[k] \in \{a, -a\}} |y[k] - s[k]| \tag{2.5}$$

and the corresponding error probability is calculated by [4]

$$P^{\text{AWGN}}[\text{error}] = Q\left(\frac{a}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{2\text{SNR}}\right). \tag{2.6}$$

In the above expression, $\text{SNR} = a^2/N_0$ is the received signal-to-noise ratio per symbol, and $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t$ is the complementary cumulative distribution function of the standard Gaussian random variable.

Now return to the communication system over a fading channel as in (2.3). Suppose that BPSK is also used. Then the detection of $s[k]$ from $y[k]$ can be done by using the following sufficient statistic:

$$r = \frac{h^*[k]}{|h[k]|} y[k] = |h[k]|s[k] + z[k] \tag{2.7}$$

where $z[k]$ has the same distribution as $w[k]$, i.e., $\mathcal{CN}(0, N_0)$. For a given channel realization $h[k]$, the error probability of detecting $s[k]$ is calculated in the same way as in the case of an AWGN channel. It is given as

$$P^{\text{Rayleigh}}[\text{error}|h[k]] = Q\left(\frac{a|h[k]|}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{2|h[k]|^2\text{SNR}}\right) \tag{2.8}$$

Since $h[k]$ is a random channel gain with distribution $\mathcal{CN}(0, 1)$, the overall error probability is computed by averaging over the random gain $h[k]$. Performing integration yields:

$$P^{\text{Rayleigh}}[\text{error}] = \text{E}\left[Q(\sqrt{2|h[k]|^2\text{SNR}})\right] = \frac{1}{2}\left(1 - \sqrt{\frac{\text{SNR}}{1+\text{SNR}}}\right) \tag{2.9}$$

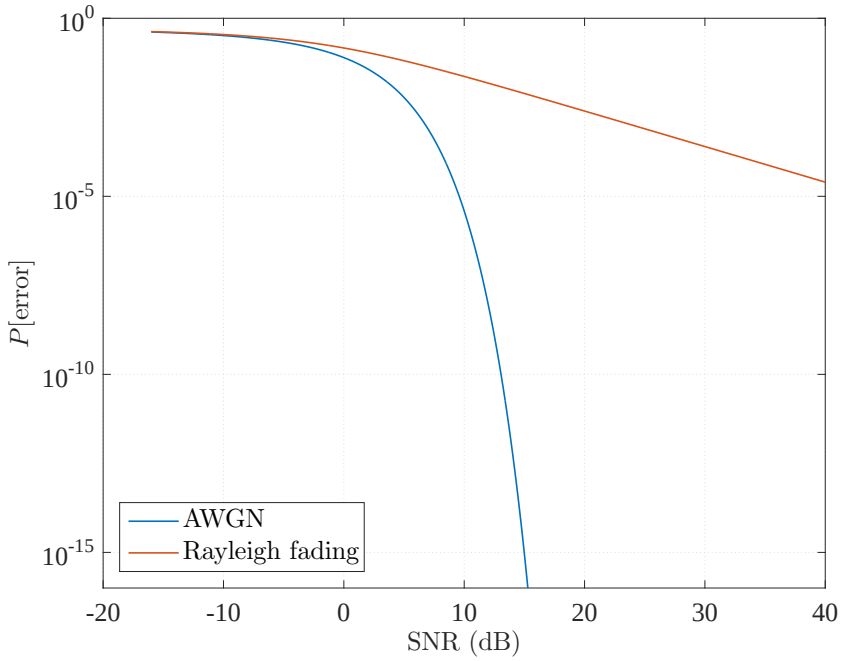

Figure 2.3 Performance of BPSK over AWGN and Rayleigh fading channels.

Figure 2.3 compares the error probabilities of BPSK over AWGN and Rayleigh fading channels. It can be seen that while the error probability over an AWGN channel decays very fast with the increasing SNR, the error probability in the case of a Rayleigh fading channel decays very slowly. At high SNR, applying Taylor series expansion yields

$$P^{\text{Rayleigh}}[\text{error}] \approx \frac{1}{4\text{SNR}} \tag{2.10}$$

The above expression shows that the diversity order, which is the power of the SNR in the denominator, is one for this system. The reason that the system over Rayleigh fading

channel performs very poorly and has the diversity order of one is that there is only one wireless link from the transmit antenna to the receive antenna. If this link suffers a "deep" fade, then the transmitted signal is detected incorrectly at the receiver. To overcome this disadvantage, diversity techniques should be used to improve the error performance. In diversity systems, many replicas of the original signal are created and transmitted to the receiver by different resources (e.g., time, frequency, space) so that the chance of having the transmitted signal experiencing a "deep" fade decreases. While diversity techniques using the time and frequency resources cost time and bandwidth expansions, the diversity technique using multiple antennas creates different spatial paths (i.e., more wireless links) from the source to the destination and it does not require time or bandwidth expansion. In the next section, the multiple-antenna wireless communication systems are described.

2.2 Multiple-Antenna Wireless Communication Systems

As discussed before, the most general form of a multiple-antenna system is known as a multiple-input multiple-output (MIMO) system in which multiple antennas are implemented at both the transmitter and receiver. Two special cases of MIMO systems are (i) single-input multiple-output (SIMO) and (ii) multiple-input single-output (MISO). As obvious from the names, a SIMO system uses multiple receive antennas to provide receive diversity, whereas a MISO system has multiple antennas at the transmitter to provide transmit diversity.

Figure 2.4 depicts a receiver equipped with N_r receive antennas to collect a signal transmitted from a single antenna. For a given time slot, the equivalent discrete-time baseband signal at the ith receive antenna can be written as

$$y_i = h_i s + w_i, \quad i = 1, \ldots, N_r \tag{2.11}$$

where s is the transmitted symbol drawn from a constellation Ψ (e.g., QPSK or 16-QAM), h_i is the channel coefficient representing the flat fading channel between the transmit antenna and the ith receive antenna, and $w_i \sim \mathcal{CN}(0, N_0)$ denotes the additive white Gaussian noise (AWGN) component. In vector form, the equivalent discrete-time baseband signal can be expressed as

$$\mathbf{y} = \mathbf{h}s + \mathbf{w} \tag{2.12}$$

where $\mathbf{y} = [y_1, \cdots, y_{N_r}]^T$, $\mathbf{h} = [h_1, \cdots, h_{N_r}]^T$ and $\mathbf{w} = [w_1, \cdots, w_{N_r}]^T$. To achieve receive diversity, the received signals from multiple antennas need to be combined properly using some combining method [18].

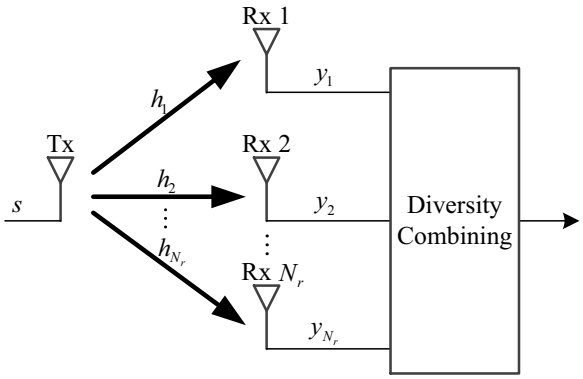

Figure 2.4 Spatial diversity provided by using multiple receive antennas.

There are two main combining methods that can be used at the receiver. The optimum combining method is the maximum-ratio combining (MRC). The MRC method weights the received signal at each receive antenna in proportion to the signal strength and also aligns the phases of the signals before summation in order to maximize the output signal-to-noise ratio (SNR) [4]. The sufficient statistic provided by the MRC is given as

$$\frac{\mathbf{h}^H}{\|\mathbf{h}\|}\mathbf{y} = \|\mathbf{h}\|\, s + \frac{\mathbf{h}^H}{\|\mathbf{h}\|}\mathbf{w} \tag{2.13}$$

This is an equivalent scalar detection problem with noise $(\mathbf{h}^H/\|\mathbf{h}\|)\mathbf{w} \sim \mathcal{CN}(0, N_0)$. For BPSK modulation, with $s = \pm a$, the error probability given the channel realization $\mathbf{h}$ can be derived as in (2.8). That is,

$$P^{\mathrm{MRC}}[\mathrm{error}|\mathbf{h}] = Q(\sqrt{2\,\|\mathbf{h}\|^2\,\mathrm{SNR}}) \tag{2.14}$$

where $\mathrm{SNR} = a^2/N_0$. Under Rayleigh fading with each gain $h_i \sim \mathcal{CN}(0,1)$, $\|\mathbf{h}\|^2 = \sum_{i=1}^{N_r} |h_i|^2$ is Chi-square distributed with $2N_r$ degrees of freedom, whose pdf is given by

$$f(x) = \frac{1}{(N_r-1)!}x^{N_r-1}\mathrm{e}^{-x}, \quad x \geq 0. \tag{2.15}$$

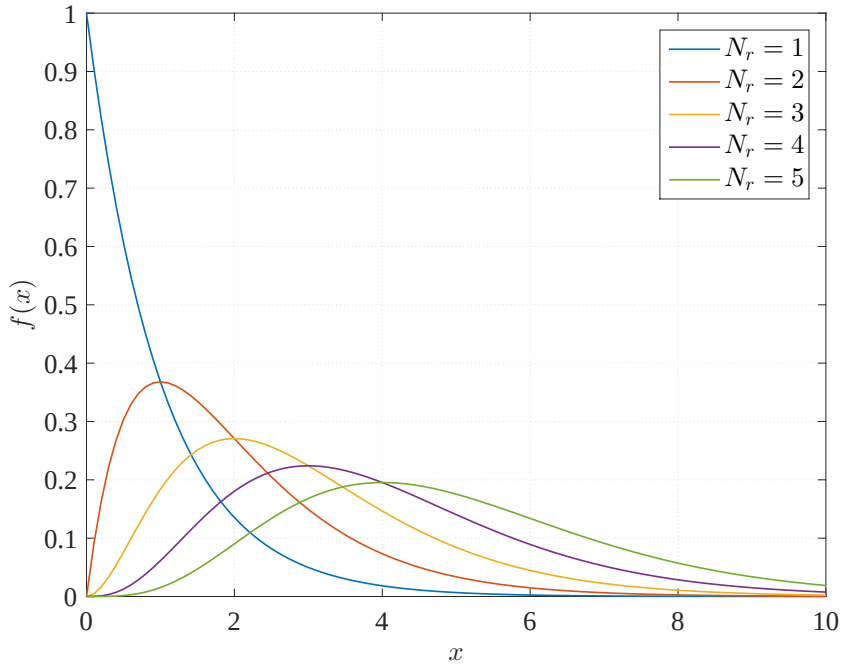

Figure 2.5 The probability density function of $\|\mathbf{h}\|^2$ for different values of N_r.

Figure 2.5 plots the distribution of $\|\mathbf{h}\|^2$ for different values of N_r. The average error probability can be computed by

$$P^{\text{MRC}}[\text{error}] = \text{E}\left[Q(\sqrt{2\,\|\mathbf{h}\|^2\,\text{SNR}})\right] = \int_0^\infty Q(\sqrt{2x\text{SNR}})f(x)\text{d}x \tag{2.16}$$
$$= \left(\frac{1-\mu}{2}\right)^{N_r} \sum_{\ell=0}^{N_r-1} \binom{N_r-1+\ell}{\ell} \left(\frac{1+\mu}{2}\right)^{\ell}$$

where

$$\mu = \sqrt{\frac{\text{SNR}}{1+\text{SNR}}} \tag{2.17}$$

The error probability as a function of SNR for different numbers of receive antenna N_r is plotted in Figure 2.6. As can be seen, increasing N_r dramatically decreases the error probability. At high SNR, using Taylor series expansion gives the approximation:

$$P^{\text{MRC}}[\text{error}] \approx \binom{2N_r-1}{N_r} \frac{1}{(4\text{SNR})^{N_r}} \tag{2.18}$$

The above expression relates the probability of error to the diversity order of the system, which is N_r.

Another important combining method is selection combining (SC). In the SC method, the received signals at all the receive antennas are compared and the one that has the highest

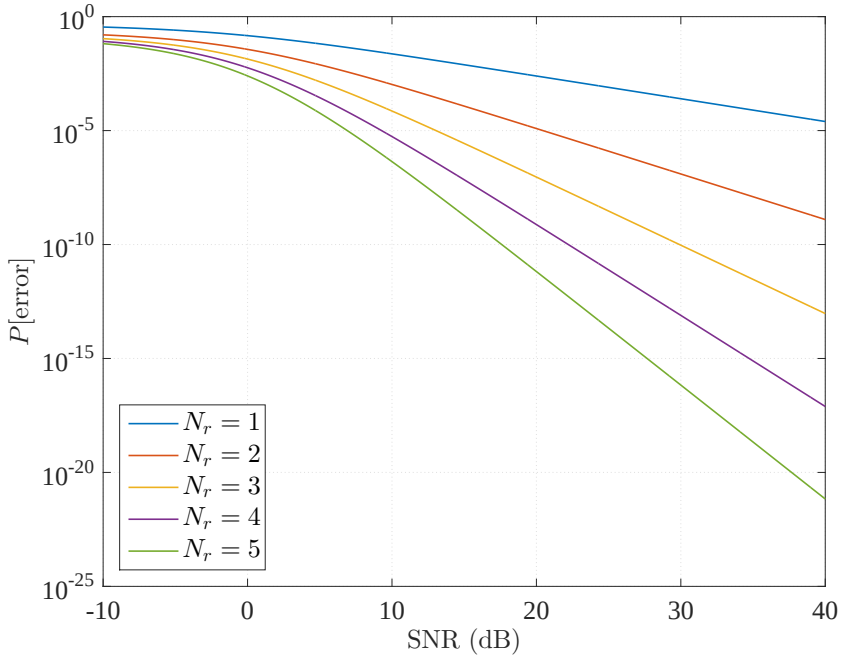

Figure 2.6 Error probability with the MRC method for different numbers of N_r.

"reliability" is selected. The output of the selection combiner is therefore y_j, where

$$j = \arg \max_{i=1,\cdots,N_r} |y_i| \tag{2.19}$$

The SC method is simpler than the MRC method, however, its performance is inferior to that of the MRC method [19].

In a multiple-input single-output (MISO) system, the transmit-diversity gain is realized by providing signal copies over N_t channels, linking N_t transmit antennas to a single receive antenna. This configuration is more suitable for applications such as in the downlink of a mobile cellular network in which the base station can be equipped with multiple antennas, while at the receiver side (mobile device) only one antenna is used (due to size limitation). Collecting the diversity in a MISO system is not as easy as in the SIMO system since the signals originating from the multiple transmit antennas are superimposed at the single receive antenna. But even more challenging is the design of a transmission scheme that can achieve the maximum available spatial diversity. A well-known technique to enable transmit diversity is space-time coding (STC), of which the Alamouti scheme is the most famous example. Alamouti presents a remarkably simple scheme to achieve transmit diversity with two transmit antennas without any loss of bandwidth efficiency [20]. Figure 2.7 depicts how

the Alamouti scheme transmits two symbols, s_1 and s_2, over two symbol intervals (time periods). In the first symbol interval, the scheme transmits s_1 from antenna 1 and s_2 from antenna 2. In the next symbol interval, symbol $-s_2^*$ is transmitted from antenna 1 and s_1^* is transmitted from antenna 2, where the superscript * represents complex conjugate operation. The transmitted codeword is thus given as

$$\mathbf{S} = \begin{pmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{pmatrix} \tag{2.20}$$

Here, it is assumed that the channel gains are quasi-static (i.e., they are constant during two time slots). Then, the received signals at the single-antenna receiver over two time slots are

$$\mathbf{y} = [y_1, y_2]^T = \mathbf{S}\mathbf{h} + \mathbf{w} \tag{2.21}$$

where $\mathbf{h} = [h_1, h_2]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_2)$ is the channel vector and $\mathbf{w} = [w_1, w_2]^T \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_2)$ is the noise vector.

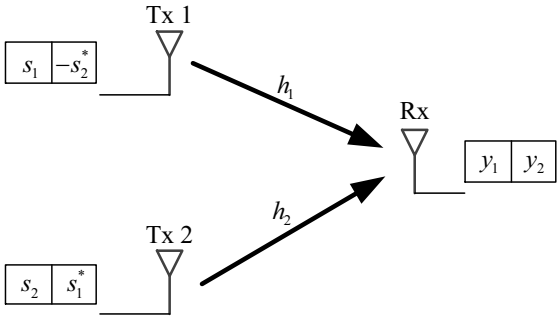

Figure 2.7 Spatial diversity provided by using multiple transmit antennas.

After some simple manipulations, (2.21) can be re-written as

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2^* \end{bmatrix} \tag{2.22}$$

The very important property that follows from the specific structure of the Alamouti STC is that the columns of the above square matrix are orthogonal, regardless of the actual values of the channel coefficients. Hence, with known channel information at the receiver, the transmitted symbols can be separately detected by projecting $[y_1, y_2^*]^T$ onto each of the two columns to obtain the sufficient statistics as follows:

$$h_1^* y_1 + h_2 y_2^* = (|h_1|^2 + |h_2|^2)s_1 + \tilde{w}_1 \tag{2.23}$$

$$h_2^* y_1 - h_1 y_2^* = (|h_1|^2 + |h_2|^2) s_2 + \tilde{w}_2 \tag{2.24}$$

where $\tilde{w}_1 = h_1^* w_1 + h_2 w_2^*$ and $\tilde{w}_2 = h_2^* w_1 - h_1 w_2^*$ are effective noise components. Based on the above expression, the information symbols s_1 and s_2 can be detected in the same way as for single-antenna system, albeit with more favorable effective channel gain of $|h_1|^2 + |h_2|^2$. In fact, such an effective channel gain yields a diversity order of two [1].

However, an orthogonal design of full-rate STC is only known for the case of two transmit antennas (Alamouti scheme), and there is no known solution for a higher number of transmit antennas. This means that, the design of STCs for more than two transmit antennas must sacrifice a portion of the data rate to achieve full orthogonality and, hence, full diversity [7]. In general, the maximum spectral efficiency of full-diversity STC systems is one symbol per symbol duration for any number of transmit antennas.

A different approach in exploiting multiple transmit antennas is called spatial multiplexing. This approach includes V-BLAST (Vertical Bell Labs Layered Space Time) and D-BLAST (Diagonal Bell Labs Layered Space Time). In a V-BLAST system [9], a high level of inter-channel interference (ICI) occurs at the receiver since all antennas transmit their own data streams at the same time. This increases the complexity of an optimal decoder, while low-complexity suboptimum linear decoders, such as the minimum mean square error (MMSE) decoder, suffer significant degradation in the system's error performance. Although a V-BLAST system can drastically increase the data rate at the cost of high computational complexity, it does not achieve transmit diversity [1]. An alternative of the V-BLAST is D-BLAST, in which the transmission is designed so that a data stream experiences different channel gains (while in V-BLAST a data stream always experiences the same channel). To illustrate the concept of D-BLAST, consider two data streams $\mathbf{a}$ and $\mathbf{b}$, and a system with two transmit antennas. Suppose that stream $\mathbf{a}$ is made of two sub-streams $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ and stream $\mathbf{b}$ consists of two sub-streams $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$. Each sub-stream can be seen as a block of symbols. The transmitted codeword $\mathbf{C}$ in D-BLAST is written as

$$\mathbf{C} = \begin{bmatrix} \mathbf{a}^{(1)} & \mathbf{b}^{(1)} & \\ & \mathbf{a}^{(2)} & \mathbf{b}^{(2)} \end{bmatrix}. \tag{2.25}$$

During the first stream period, only $\mathbf{a}^{(1)}$ is transmitted on the first antenna, while the second

antenna sends nothing. In the second stream period, $\mathbf{b}^{(1)}$ is sent on antenna 1, while $\mathbf{a}^{(2)}$ is sent on antenna 2. In the last stream period, $\mathbf{b}^{(2)}$ is sent on antenna 2, while antenna 1 sends nothing. In V-BLAST, there is no coding across sub-channels, error therefore occurs whenever one of the sub-channels is in a deep fade. On the other hand, by coding across sub-channels, D-BLAST can average the randomness of the individual sub-channels and improve the performance over V-BLAST [4].

As pointed out before, SIMO and MISO systems are special cases of MIMO systems. Generally, a transmit/receive diversity MIMO system can be realized with multiple antennas at both the transmitter and receiver to achieve the transmit and receive diversity gains simultaneously. With N_t transmit antennas and N_r receive antennas, the maximum available diversity order of a MIMO system is $N_t \times N_r$ [7].

2.3 Spatial Modulation

The previous section has discussed MIMO systems with multiple antennas at the transmitter and receiver. In this section, spatial modulation (SM) is introduced as a special form of MIMO. The key idea of SM is to use antenna index as an additional dimension to convey information bits. The spectral efficiency of SM is

$$\lfloor \log_2(N_t) \rfloor + \log_2 M, \quad \text{bits/s/Hz} \tag{2.26}$$

where N_t is the number of transmit antennas and M is the size of the employed constellation. Usually, N_t should be a power of two.

To clearly understand the operation of SM, consider an example of SM-MIMO with $N_t = 4$ and QPSK constellation as illustrated in Figure 2.8. In this example, the system transmits 4 bits per symbol duration. The first and second blocks of 4 bits are assumed to be 0010 and 1101. In the first block, the first two bits 00 are used to pick the green antenna, while the last two bits 10 are mapped to symbol $-1 + j$ of the QPSK constellation. The transmit symbol vector over the four transmit antennas is then $[0\ 0\ 0\ \ -1 + j]^T$. Similarly, in the second symbol duration, the first two bits 11 picks the red antenna while the last two bits 01 selects the QPSK symbol $1 - j$ and the transmit symbol vector in this duration is

$[1-j\ 0\ 0\ 0]^T$.

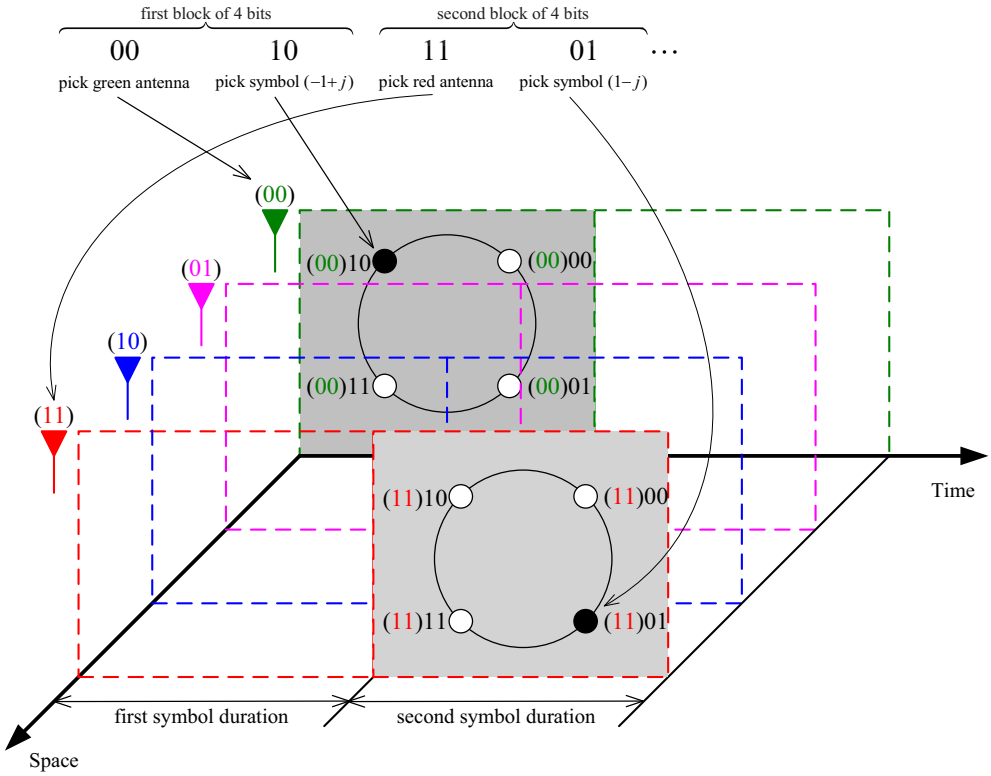

Figure 2.8 Example of SM-MIMO with $N_t = 4$ transmit antennas and QPSK constellation [1].

SM Transmitter

A general block diagram of spatial modulation is shown in Figure 2.9. Here $\mathbf{q}(k)$ is a vector of n information bits to be transmitted over the kth symbol duration. The binary vector is mapped into a column vector $\mathbf{x}(k)$ of size N_t such that only one element in the resulting vector is different from zero. Suppose that the ith element in column vector $\mathbf{x}(k)$ is nonzero and indicated as x_ℓ, where $\ell \in [1, \cdots, M]$ and M is the constellation size. The symbol x_ℓ is then transmitted from the ith transmit antenna over the MIMO channel $\mathbf{H}(k)$. The MIMO channel matrix $\mathbf{H}(k)$ can be written as a set of vectors where each vector corresponds to the channel path gains between one transmit antenna to N_r receive antennas as follows:

$$\mathbf{H}(k) = [\mathbf{h}_1\ \mathbf{h}_2 \cdots \mathbf{h}_{N_t}] \tag{2.27}$$

where

$$\mathbf{h}_\nu = [h_{1,\nu}\ h_{2,\nu} \cdots h_{N_r,\nu}]^T, \quad \nu = 1, \ldots, N_t. \tag{2.28}$$

It is pointed out that, for simplicity of presentation, the time index k is omitted in the channel vectors. The receive signal vector is then given by

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{w}(k) = \mathbf{h}_i x_\ell + \mathbf{w}(k) \tag{2.29}$$

where $\mathbf{w}(k)$ is the additive white Gaussian noise vector.

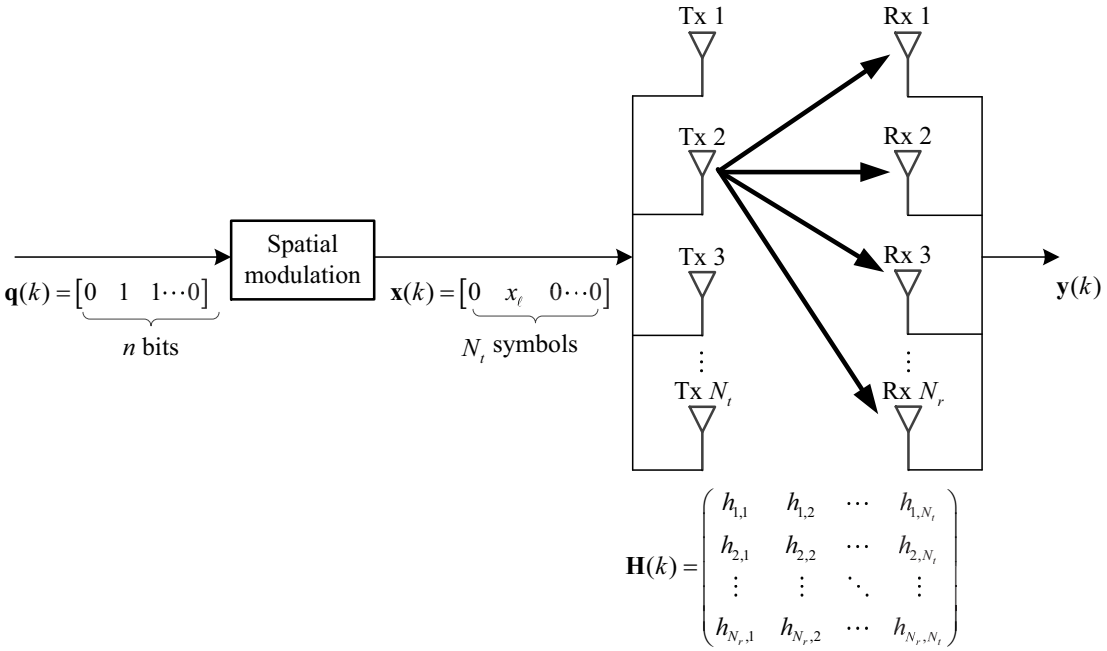

Figure 2.9 General block diagram of spatial modulation.

SM Receiver

Similar to conventional MIMO systems, different detection approaches can be developed for SM systems. For example, an SM receiver can implement combining methods such as the MRC or SC to extract the receive-diversity gain. The major difference between the detection processes in SM and conventional MIMO systems is that the SM receiver needs to detect both the antenna index and constellation symbol, while a receiver for a MIMO system needs to detect only the transmitted constellation symbol. The detection rule of the ML receiver for SM can be expressed as

$$[\tilde{i}, \tilde{\ell}] = \arg \min_{1 \le i \le N_t,\ 1 \le \ell \le M} \|\mathbf{y}(k) - \mathbf{h}_i x_\ell\|. \tag{2.30}$$

In the following, the advantages and disadvantages of SM systems are discussed, which are largely based on [21].

Advantages of SM-MIMO

- *Higher throughput*: By using the additional dimension (i.e., antenna index), the spectral efficiency of SM-MIMO is higher than that of single-antenna and OSTBC-MIMO transmission. This improved spectrum efficiency translates into the reduction of the RF output power [22].

- *Simpler receiver design*: Since only a single transmit antenna is active in every channel use, SM-MIMO is not affected by ICI, hence the optimum performance can be obtained at a single-stream decoding complexity [23].

- *Simpler transmitter design*: By activating only one transmit antenna, SM-MIMO can be implemented with a single RF chain, which is inexpensive and easy to employ [24]. Thus, the implementation of multiple expensive and bulky power amplifiers, RF filters/mixers, analog-to-digital converters, and RF coaxial cables can be avoided [10].

- *Lower transmit power supply*: Since the multiplexing gain is achieved by a single-RF source, SM-MIMO reduces the total consumed power required for the same RF output power. This results in an energy efficiency gain [25].

- *Better efficiency of the power amplifiers*: The power efficiency of power amplifiers decreases when increasing the linearity requirements of the modulation scheme (e.g., QAM) [26]. The reduced linearity requirements of constant-envelope modulation, which is recently shown as a suitable modulation for SM-MIMO, increase the efficiency of power amplifiers, which, in turn, reduce the total power consumption of the transmitters [26].

These important advantages of SM-MIMO, however, come with some fundamental trade-offs. The main disadvantages of SM-MIMO are discussed next.

Disadvantages of SM-MIMO

- *Spectral efficiency suboptimality*: SM-MIMO offers a lower throughput than the spatial-multiplexing MIMO (SMUX-MIMO), such as V-BLAST. In particular, the spectral efficiency of SMUX-MIMO increases linearly with N_t, while that of SM-MIMO increases logarithmically with N_t. This disadvantage may, however, be offset by the emerging large-scale MIMO and millimeter wave communications, which could employ hundreds of antennas at the transmitter and receiver [10, 27, 28].

- *Fast antenna switching*: The active transmit antenna in SM-MIMO is changed in each transmission time. As a consequence, a single-RF implementation needs a sufficiently fast RF switch operation. Fortunately, high speed RF switches are available for a wide range of frequency bands [29–31].

- *Time-limited pulse shaping*: The transmitted signals are usually designed to have a flat spectrum and a fast roll-off by using appropriate shaping filters before transmission. Commonly adopted shaping filters satisfy the Nyquist criterion, thus they are bandwidth-limited and hence have long (theoretically, infinite) impulse response. On the other hand, SM-MIMO based on a symbol-time switching mechanism is better suited for time-limited pulse shapes [32], which results in a bandwidth expansion. Thus, in SM-MIMO systems, a good trade-off between a limited time duration and a practical bandwidth needs to be considered [33].

- *Training overhead*: The SM-MIMO receiver needs to estimate the channel gains of all the transmit antennas for the ML detection. A large scale SM-MIMO implementation may require a non-negligible training overhead for channel estimation. Moreover, a single active antenna in SM-MIMO may increase further the training overhead, since the channel gains of all the transmit antennas cannot be estimated simultaneously [34].

- *Directional beamforming*: Large scale SM-MIMO systems equipped with tens or hundreds of antennas may be suitable for millimeter-wave frequencies. In this frequency band, however, directional beamforming is necessary to overcome the higher pathloss and the higher noise level compared to the current frequency bands. Thus, SM-MIMO systems need to provide a directional beamforming gain.

2.4 OFDM and Spatial Modulation-OFDM

The previous section discusses spatial modulation under frequency-flat fading channels. However, for many applications in current and future generations of wireless communication networks, the wireless channel is highly frequency selective, which results in inter-symbol interference (ISI) at the receiver. If not properly handled, ISI can severely deteriorate the system error probability. For a frequency-selective fading channel, orthogonal-frequency division multiplexing (OFDM) is a popular technique to combat ISI and exploit frequency diversity. As such, combining SM and OFDM is promising to deal with frequency selective channels. The following description of OFDM largely follows [1].

2.4.1 OFDM

For a wideband channel, the frequency response $H(f)$ typically exhibits a noticeable variation over the frequency, i.e., the channel is *frequency-selective*. The degree of frequency selectivity of a channel is usually measured by the so-called channel's *coherence bandwidth*, denoted by B_c. Loosely speaking, B_c is the frequency duration over which the channel frequency response is approximately constant. A related parameter that characterizes the frequency-selectivity of a channel is the multipath *delay spread*, which is loosely defined as the inverse of the coherence bandwidth, i.e., $T_m = 1/B_c$. The meaning of this parameter is that, if a short transmitted pulse of duration T is transmitted over the channel, then the received signal will have a duration of approximately $T + T_m$. If T_m is comparable to T, then the consequence of *delay spread* is to cause *inter-symbol interference* (ISI) of the transmitted signal at the channel output [2].

OFDM uses multiple orthogonal carrier frequencies to send information bits. Specifically, in OFDM all the carriers can be activated all the time and QAM symbols are sent over each carrier. For each QAM subchannel in OFDM, its bandwidth is $B_N = \frac{1}{T_N}$, where T_N is the duration of each OFDM symbol. Thus, if B_N is made to be smaller than B_c (by using as many number of carriers as required), then each QAM subchannel experiences a frequency-flat fading channel. The condition $B_N < B_c$ implies that $T_N > T_m$, thus ISI can be made negligible. Flat fading and zero-ISI are desirable and this is achieved with OFDM

transmission by using a large number of subcarriers [2].

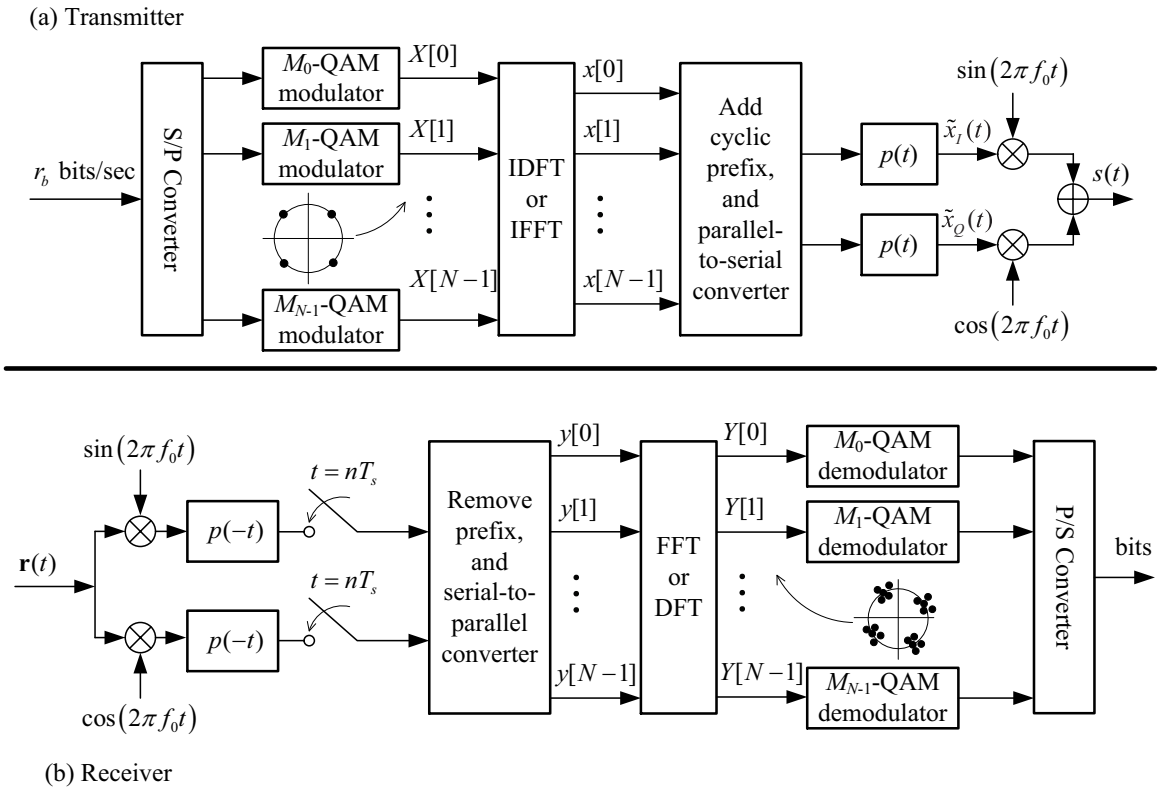

Figure 2.10 Implementation of OFDM transmitter and receiver with IDFT/DFT [2].

The implementation of OFDM transmitter and receiver uses N subcarriers corresponding to N QAM subchannels is shown in Figure 2.10 [2]. The information bit stream is processed by a QAM mapper, resulting in a complex symbol stream $X[0], X[1], \cdots, X[N-1]$. This symbol stream is passed through a serial-to-parallel converter, whose output is a set of N parallel QAM symbols $X[0], \cdots, X[N-1]$, each symbol is transmitted over each of the N subcarriers. Thus, the N symbols from the serial-to-parallel converter correspond to the discrete frequency components of the OFDM signal $s(t)$. In order to generate $s(t)$, these frequency components are converted into time samples by the inverse DFT (which could be efficiently implemented using the IFFT algorithm if N is a power of 2). The IFFT yields an *OFDM symbol* consisting of the complex-valued sequence $x[0], \cdots, x[N-1]$ of length N,

where

$$x[n] = \frac{1}{N}\sum_{i=0}^{N-1} X[i]\mathrm{e}^{j\frac{2\pi ni}{N}}, \quad 0 \leq n \leq N-1. \tag{2.31}$$

Next, a *cyclic prefix* (CP) extension is performed on the OFDM symbol to produce time samples. After CP extension, these time samples are ordered by the parallel-to-serial converter and passed through a shaping filter, resulting in the complex *baseband* OFDM signal $\tilde{x}(t) = \tilde{x}_I(t) + j\tilde{x}_Q(t)$. The baseband OFDM signal is then up-converted to frequency f_0. The transmitted signal is filter by the channel impulse response $h(t)$ and corrupted by AWGN $w(t)$. The received signal is $r(t) = s(t) * h(t) + w(t)$, where $*$ is the convolution operator. This signal is down-converted to baseband and filtered to remove the high frequency components. The output of the matched filter is sampled every T_s to obtain $y[n]$, where $T_s = T_N/N$ is the duration between two adjacent time samples.

By the operation of the CP insertion, the samples of the channel output in the range $0 \leq n \leq N-1$ can be expressed as [1, 4]:

$$y[n] = x[n] \otimes h[n], \quad 0 \leq n \leq N-1 \tag{2.32}$$

where $h[n]$ is the equivalent discrete-time channel and $\otimes$ is a circular convolution. Taking into account AWGN at the channel input, the DFT of the channel output yields

$$Y[i] = \mathrm{DFT}\{y[n] = (x[n] + w[n]) \otimes h[n]\} = X[i]H[i] + W[i] \tag{2.33}$$

where $H[i] = H(f_i)$ is the flat-fading channel gain associated with the ith subchannel and $W[i]$ is Gaussian noise component. The FFT output is parallel-to-serial converted and passed though a QAM demodulator to recover the original data. The demodulator needs the channel gains $H[i], i = 0, \cdots, N-1$ to recover the original QAM symbols by dividing out these gains: $X[i] = Y[i]/H[i]$. This process is called *frequency equalization*.

2.4.2 Spatial Modulation with OFDM (SM-OFDM)

In the previous subsection, OFDM is presented as an effective transmission method for frequency selective fading channels. This subsection discusses a combination of SM and OFDM, called SM-OFDM, that was proposed in [13].

For SM-OFDM, the symbols at the output of the spatial modulation block is grouped into N_t vectors, where N_t is the number of transmit antennas. Each vector corresponds to one of the transmit antennas. For example, all the subcarriers that should be transmitted from the first antenna are included in the first vector and all other subcarriers are set to zero. This is done by grouping the symbols at the output of the spatial modulator in vectors corresponding to their assigned transmit antenna number and setting all other symbols in that vector to zero. The OFDM modulator is applied to each vector, thus, resulting in N_t OFDM blocks to be transmitted simultaneously from the N_t transmit antennas. However, at each instant of time and for each subcarrier only one transmit antenna will be active and all other transmit antennas are off. As a result, ICI is completely avoided at the receiver input.

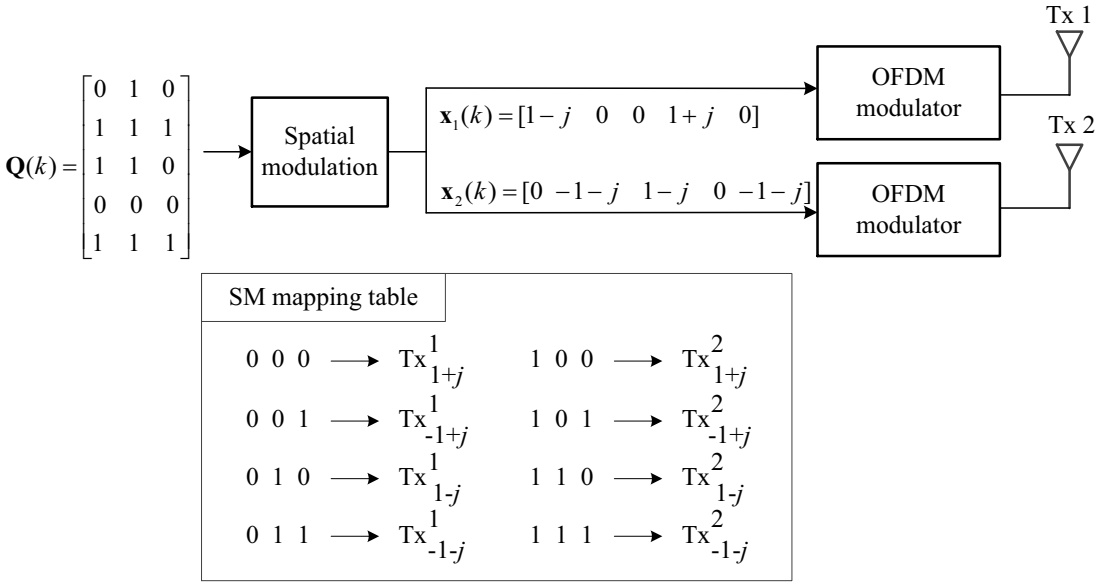

Figure 2.11 Illustration of SM-OFDM transmission.

The operation of SM-OFDM is described in more detail in Figure 2.11. Here $\mathbf{Q}(k)$ is an $N \times m$ binary matrix to be transmitted, where N is the number of OFDM subcarriers. This matrix is mapped into another matrix $\mathbf{X}(k)$ of size $N_t \times M$. The mapping operation works as follows. First, the $\mathbf{Q}(k)$ matrix is mapped into a vector containing N M-QAM symbols. Each symbol in the resulting vector corresponds to a single transmit antenna. For example, as seen in Figure 2.11, an initial three-bit input sequence of 010 is mapped

to Tx^1_{1-j}, where Tx^1_{1-j} means that the QPSK symbol $1-j$ is transmitted from the first transmit antenna. The second antenna transmits zero power on that respective sub carrier. Similarly, the second three-bit block of 111 is mapped to Tx^2_{-1-j} and so forth. Finally, arranging the resultant symbol vectors produces the $\mathbf{X}(k)$ matrix, where each row $\mathbf{x}_\lambda(k)$ contains the symbols to be transmitted from transmit antenna λ. All the other symbols that do not belong to this antenna are set to zero as seen in Figure 2.11. Afterwards, each row vector $\mathbf{x}_\lambda(k)$ is modulated using an OFDM modulator.

2.4.3 OFDM-Index Modulation (OFDM-IM)

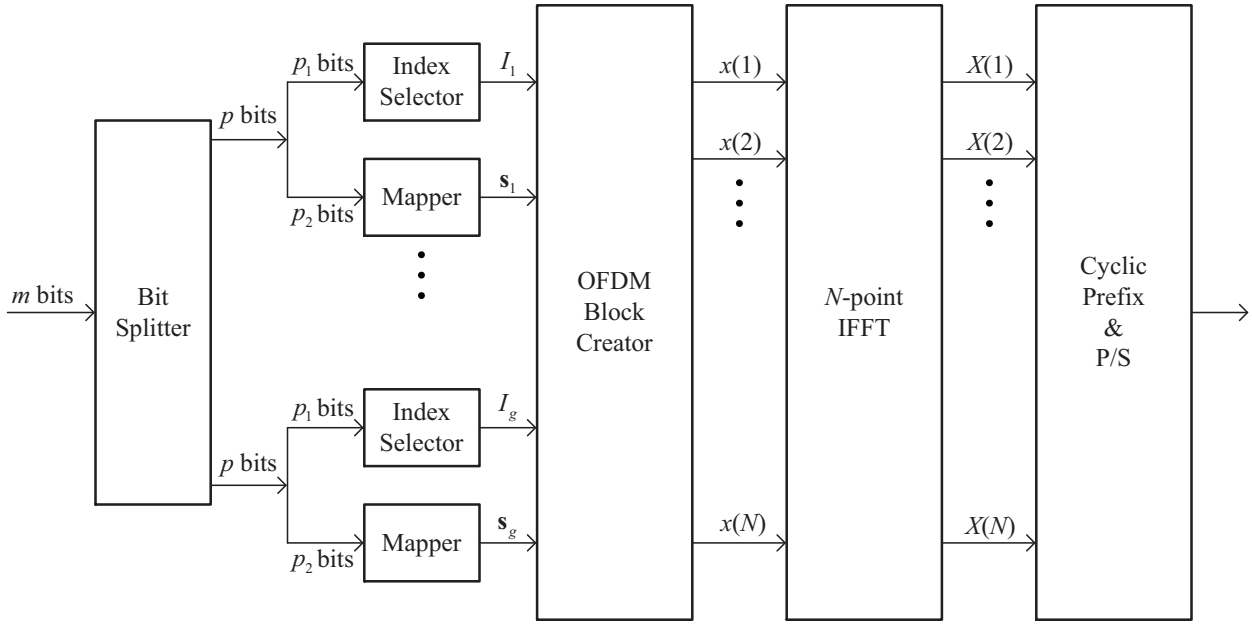

Figure 2.12 Block diagram of OFDM-IM transmitter.

Motivated by the conventional SM technique, where antenna index is used to convey information bits, recently, a novel OFDM system, called OFDM with index modulation (OFDM-IM), is proposed [35] for operating over frequency-selective fading channels. In OFDM-IM, the subcarrier index is used as an additional means to convey information bits. Specifically, the information is conveyed not only by an M-ary signal constellation as in classical OFDM, but also by the indices of the subcarriers, which are activated according to the incoming bit stream.

The operation of the OFDM-IM scheme is described with the help of a block diagram

in Figure 2.12. A total of m information bits enter the OFDM-IM transmitter for the transmission of one OFDM block. These m bits are then split into g groups each containing p bits, i.e., $m = p \cdot g$. Each group of p bits is mapped to an OFDM subblock of length n, where $n = N/g$ and N is the number of OFDM subcarriers, i.e., the size of the fast Fourier transform (FFT). Unlike the classical OFDM, this mapping operation is not only performed by means of the modulated symbols, but also by the indices of the subcarriers. Inspired by the SM concept, additional information bits are transmitted by a subset of the OFDM subcarrier indices. For a specific subblock β, only k out of n available indices are active for this purpose and they are determined by the index selector, based on the first p_1 bits of the incoming p-bit sequence. The value of p_1 is calculated by

$$p_1 = \left\lfloor \log_2 \binom{n}{k} \right\rfloor \tag{2.34}$$

At the output of the index selector, the selected indices are expressed as

$$I_\beta = \{i_{\beta,1}, \cdots, i_{\beta,k}\} \tag{2.35}$$

where $i_{\beta,\gamma} \in [1, \cdots, n]$ and $\gamma = 1, \cdots, k$. For inactive subcarriers in subblock β, the data is not transmitted. The remaining $p_2 = k \log_2 M$ bits of this sequence are mapped onto the M-ary signal constellation to determine the data symbols that modulate the subcarriers having active indices. The vector of the modulated symbols at the output of the M-ary mapper (modulator), which carries p_2 bits, is given by

$$\mathbf{s}_\beta = [\mathbf{s}_\beta(1), \mathbf{s}_\beta(1), \cdots, \mathbf{s}_\beta(k)] \tag{2.36}$$

Without loss of generality, assume that $\mathrm{E}\{\mathbf{s}_\beta \mathbf{s}_\beta^H\} = k$, i.e., the signal constellation is normalized to have unit average power. The OFDM block creator creates all of the subblocks by taking into account I_β and $\mathbf{s}_\beta$ for all β first and it then forms the $N \times 1$ main OFDM block:

$$\mathbf{x}_F = [x(1), x(2), \cdots, x(N)] \tag{2.37}$$

Unlike the classical OFDM, in the OFDM-IM scheme the vector $\mathbf{x}_F$ contains some zero terms. After this operation, the same procedure as in the classical OFDM system is applied. Specifically, the OFDM block is processed by the inverse FFT (IFFT) to form the time-domain OFDM block. At the output of the IFFT, a cyclic prefix is appended to the beginning

of the OFDM block. After parallel to serial (P/S) and digital-to-analog conversion, the signal is sent through a frequency-selective Rayleigh fading channel.

In fact, the OFDM-IM scheme benefits from the frequency selectivity of the channel by exploiting subcarrier indices as a resource for information transmission. Therefore, the error performance of the OFDM-IM scheme is significantly better than that of the classical OFDM due to the higher diversity order attained for the bits transmitted in the spatial domain of the OFDM block.

2.5 Summary

This chapter provides background relevant to the works performed in this Ph.D. research. First, basic point-to-point wireless communication systems using single antenna and multiple antennas were introduced. It was then followed by the introduction of spatial modulation systems under both frequency flat and frequency selective fading channels. For single-antenna systems, the structures of the transmitter, channel fading model and receiver were described. For multiple-antenna systems, receive diversity using two combining methods, namely MRC and SC, and transmit diversity using space-time coding and spatial multiplexing with V-BLAST and D-BLAST were presented. For the SM system operating over a flat fading channel, the transmitter and the receiver were discussed together with advantages and disadvantages of this system. Finally, for the SM system operating over a frequency selective fading channel, a combination of SM and OFDM was presented. The chapter concludes by introducing an OFDM scheme where subcarrier index is used to convey information bits. In the next chapter, novel SM schemes for frequency-flat fading channels will be introduced.

3. Spatial Modulation for Frequency-Flat Fading Channels

The major disadvantage of the conventional SM is its limit of spectral efficiency, which is only $\lfloor \log_2(N_t) \rfloor + \log_2 M$. This is low compared to $N_t \log_2 M$ achieved by spatial multiplexing. In this chapter, novel transmission techniques based on SM are developed to improve spectral efficiency of the conventional SM over frequency-flat fading channels.

First, a transmission scheme is proposed for two active antennas. Although, having two active antennas with two RF chains is more complicated than activating a single antenna in the conventional SM, the benefit is that the spectral efficiency is enhanced and the technique of space-time block coding can be applied to achieve a transmit diversity order of two.

Another transmission scheme is developed for space-shift keying (SSK) which is the simplified form of the conventional SM. In SSK, only antenna index is used to convey the information bits while amplitude/phase modulation is removed. This makes SSK enjoying very low decoding complexity. In this second new SM-based transmission scheme, the spectral efficiency is improved by multiplexing two orthogonal channels: one in-phase channel and one quadrature channel.

The third transmission scheme is proposed to generalize the second scheme so that amplitude/phase modulation can be allowed. This third scheme also makes use of the in-phase and the quadrature components to improve the spectral efficiency. Specifically, the real and imaginary parts of two constellation symbols are transmitted on two different layers of antennas.

As a final contribution of this chapter, an optimal constellation design for the quadrature

SM technique is investigated which minimizes the system's error probability. In quadrature SM, the number of bits sent by antenna index is twice that in the conventional SM.

3.1 High-Rate Space-Time Block Coded Spatial Modulation

Focusing on the case that two antennas are active among available transmitted antennas, this part of the research proposes a SM technique that is better than the state-of-the-art schemes introduced by Basar et al. in [36] and Wang in [37]. The case of having two active antennas (i.e., two RF chains) is of great practical interest since it is only slightly more complex than the original SM scheme while it offers both increased spatial diversity as well as higher transmission rate. The scheme proposed in [36], called space-time block coded spatial modulation (STBC-SM), makes use of the famous Alamouti STBC as a core. In contrast, our proposed scheme[1] can increase the data rate and achieve a transmit diversity order of two by making use of the high-rate STBC in [39]. Alamouti code and high-rate STBC code is shown in Figure 3.1. To distinguish it from the STBC-SM scheme in [36], the scheme proposed here shall be referred to as high-rate space-time block coded spatial modulation (HR-STBC-SM). In addition to the coding gain analysis of the proposed HR-STBC-SM scheme, a simplified ML detection is also developed. Simulation results shall demonstrate that the HR-STBC-SM scheme outperforms the STBC-SM scheme at high spectral efficiency. It also outperforms the scheme recently proposed in [37] that is based on an error-correcting code.

3.1.1 Proposed HR-STBC-SM Scheme

Recall that the rate of the Alamouti STBC is one symbol per one time slot, i.e., 1 symbol per channel use (pcu). In contrast, the high-rate STBC proposed in [39] transmits two symbols over one time slot, i.e., its rate is 2 symbols pcu. The transmission matrix of such a high-rate code is as follows:

$$X(x_1, x_2, x_3, x_4) = \begin{pmatrix} ax_1 + bx_3 & ax_2 + bx_4 \\ -cx_2^* - dx_4^* & cx_1^* + dx_3^* \end{pmatrix}, \tag{3.1}$$

[1]The contributions in this section are presented in [38].

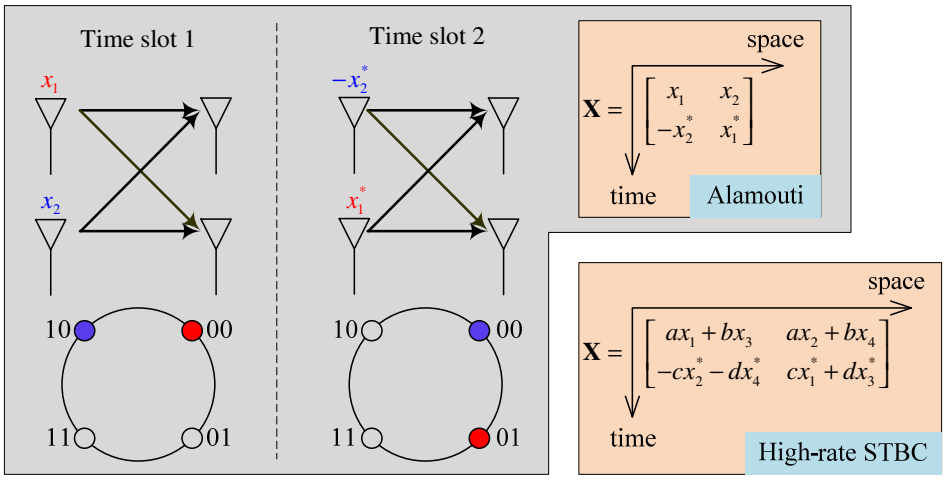

Figure 3.1 Alamouti and high-rate space-time Codes.

where $\{x_i\}_{i=1}^4$ are information symbols belonging to a standard M-ary constellation Ψ. The rows of the above 2×2 matrix correspond to the symbol times, while the columns correspond to the transmit antennas. In fact, this high-rate code is constructed as a linear combination of two Alamouti space-time matrices and the parameters a, b, c and d can be optimized to maximize the minimum coding gain. It was shown in [39] that $a = \frac{1}{\sqrt{2}}$, $b = \frac{(1-\sqrt{7})+i(1+\sqrt{7})}{4\sqrt{2}}$, $c = \frac{1}{\sqrt{2}}$ and $d = -ib$ are the optimal values. This high-rate code is chosen to replace the Alamouti code in the construction of the STBC-SM scheme because it achieves a higher coding gain than the Alamouti code for the same transmission rate measured in bits pcu, i.e., bits/s/Hz. This is because for the same transmission rate in bits/s/Hz, the constellation used in the high-rate code can have a lower order when compared to the constellation used in the Alamouti code.

In the following, the operation of the proposed HR-STBC-SM scheme is described with an example of 4 available transmit antennas. With 4 available transmit antennas, the maximum number of different antenna pairs is $\binom{4}{2} = 6$. This means that only 2 bits can be used to index 4 antenna pairs. The high-rate code is applied for the 4 selected antennas pairs as follows:

$$\mathbf{X}_1(x_1, x_2, x_3, x_4) = \begin{pmatrix} ax_1 + bx_3 & ax_2 + bx_4 & 0 & 0 \\ -cx_2^* - dx_4^* & cx_1^* + dx_3^* & 0 & 0 \end{pmatrix}$$

$$\mathbf{X}_2(x_1,x_2,x_3,x_4)=\begin{pmatrix}0 & 0 & ax_1+bx_3 & ax_2+bx_4\\ 0 & 0 & -cx_2^*-dx_4^* & cx_1^*+dx_3^*\end{pmatrix}$$

$$\mathbf{X}_3(x_1,x_2,x_3,x_4)=\begin{pmatrix}0 & ax_1+bx_3 & ax_2+bx_4 & 0\\ 0 & -cx_2^*-dx_4^* & cx_1^*+dx_3^* & 0\end{pmatrix}\mathrm{e}^{j\phi}$$

$$\mathbf{X}_4(x_1,x_2,x_3,x_4)=\begin{pmatrix}ax_2+bx_4 & 0 & 0 & ax_1+bx_3\\ cx_1^*+dx_3^* & 0 & 0 & -cx_2^*-dx_4^*\end{pmatrix}\mathrm{e}^{j\phi}$$

As can be seen from the above 2×4 matrices, there are only two non-zero columns, which guarantees that only two antennas are active at each transmission time. The high-rate code itself conveys 4 information symbols for each 2 time slots and these symbols are drawn from M-ary constellation Ψ.

If the same constellation is used in both the HR-STBC-SM and STBC-SM schemes, then the rate of the former is always higher than the rate of the latter. For example, if the constellation is QPSK, the spectral efficiency of the HR-STBC-SM scheme is 5 bits/s/Hz, while that of the STBC-SM is only 3 bits/s/Hz. The above four transmission matrices are grouped into two different codebooks Ω_1 and Ω_2 as $\Theta=\{(\mathbf{X}_1,\mathbf{X}_2)\in\Omega_1,(\mathbf{X}_3,\mathbf{X}_4)\in\Omega_2\}$. A rotation is applied for codewords in Ω_2 in order to preserve the diversity gain of the system. If such a rotation is not implemented, the difference matrix between $\mathbf{X}_1$ and $\mathbf{X}_3$ will not be a full rank, which reduces the diversity gain. The rotation angle ϕ needs to be optimized to maximize the coding gain. For QPSK with $E\{|x_i|^2\}=1$, the optimal angle ϕ is found to be 1.13 radian and the corresponding minimum coding gain is 0.1846, where the minimum coding gain is defined as

$$\Delta=\min_{\substack{\mathbf{X}_i,\mathbf{X}_j\in\Theta\\ \mathbf{X}_i\neq\mathbf{X}_j}}\det(\mathbf{X}_i-\mathbf{X}_j)(\mathbf{X}_i-\mathbf{X}_j)^H \tag{3.2}$$

Similar to [36], the general framework of the proposed HR-STBC-SM scheme for an arbitrary number of transmit antennas is described as follows:

1. Determine the number of codewords in each codebook as $n=\lfloor\frac{N_t}{2}\rfloor$, where N_t is the number of available transmit antenna.

2. Determine the total number of codewords as $q = \lfloor \binom{N_t}{2} \rfloor_{2^p}$.

3. Determine the number of codebooks as $\lceil \frac{q}{n} \rceil$. The number of codebooks is also the number of rotation angles that need to be optimized in order to maximize the minimum coding gain. The larger the number of needed rotation angles is, the smaller the minimum coding gain becomes.

Given the number of codewords q, the spectral efficiency of the HR-STBC-SM scheme is $m = \frac{1}{2}\log_2 q + 2\log_2 M$ (bits/s/Hz).

Table 3.1 shows the minimum coding gains and optimized angles for various numbers of available transmit antennas. In calculating the minimum coding gains, both BPSK and QPSK constellations are normalized to have unit average energy.

Table 3.1 Minimum coding gains and optimized angles for the cases of 4, 6 and 8 available transmit antennas.

N_t	BPSK	Angles	QPSK	Angles
4	1.5	ϕ=1.7	0.1846	ϕ=1.13
6	1	$\phi_2 = \frac{\pi}{3}$	0.1497	$\phi_2 = \frac{\pi}{6}$
		$\phi_3 = \frac{2\pi}{3}$		$\phi_3 = \frac{\pi}{3}$
8	0.5858	$\phi_2 = \frac{\pi}{4}$	0.1015	$\phi_2 = \frac{\pi}{8}$
		$\phi_3 = \frac{\pi}{2}$		$\phi_3 = \frac{\pi}{4}$
		$\phi_4 = \frac{3\pi}{4}$		$\phi_4 = \frac{3\pi}{8}$

3.1.2 Low-Complexity ML Detection Algorithm

Let $\mathbf{H}$ be a $N_t \times n_R$ channel gain matrix corresponding to a flat-fading MIMO system with N_t transmit and n_R receive antennas. For Rayleigh fading, the entries of $\mathbf{H}$ are modelled as independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and unit variance. It is further assumed that the fading is such that $\mathbf{H}$ varies independently from one codeword to another and is invariant during the transmission of a codeword, i.e., block fading. The channel matrix $\mathbf{H}$ is perfectly estimated at the receiver,

but unknown at the transmitter. With $\mathbf{X} \in \Theta$ being the $2 \times N_t$ HR-STBC-SM transmission matrix, the $2 \times n_R$ received signal matrix $\mathbf{Y}$ is given as

$$\mathbf{Y} = \sqrt{\frac{\rho}{\mu}}\mathbf{X}\mathbf{H} + \mathbf{N}, \tag{3.3}$$

where μ is a normalization factor to ensure that ρ is the average SNR at each receive antenna, $\mathbf{N}$ is a $2 \times n_R$ matrix representing AWGN, whose elements are i.i.d complex Gaussian random variables with zero mean and unit variance.

The ML detection chooses a codeword that minimizes the following decision metric:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X} \in \Theta} \left\| \mathbf{Y} - \sqrt{\frac{\rho}{\mu}}\mathbf{X}\mathbf{H} \right\|^2 \tag{3.4}$$

Since the HR-STBC-SM transmission matrix $\mathbf{X}$ contains 4 information symbols, the ML detection needs to search over qM^4 candidates to find the minimum of the above metric.

To reduce the computational complexity of the ML detection, (3.3) can be rewritten in the following form:

$$\mathbf{y} = \sqrt{\frac{\rho}{\mu}}\mathbb{H}_\ell \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \mathbf{n}, \tag{3.5}$$

where $\mathbf{y}$ and $\mathbf{n}$ are $2n_R$-length column vectors obtained by vectorizing matrices $\mathbf{Y}$ and $\mathbf{N}$ as

$$\mathbf{y} \triangleq [\mathbf{Y}(1,1), \ldots, \mathbf{Y}(1,n_R), \mathbf{Y}^*(2,1), \ldots, \mathbf{Y}^*(2,n_R)]^T \tag{3.6}$$

$$\mathbf{n} \triangleq [\mathbf{N}(1,1), \ldots, \mathbf{N}(1,n_R), \mathbf{N}^*(2,1), \ldots, \mathbf{N}^*(2,n_R)]^T \tag{3.7}$$

In (3.5), $\mathbb{H}_\ell$ is the $2n_R \times 4$ equivalent channel matrix corresponding to the transmitted codeword $\mathbf{X}_\ell$, $\ell = 1, 2, \cdots, q$. An example of 4 equivalent channel matrices for the case of

$N_t = 4$ is as follows:

$$\mathbb{H}_1 = \begin{pmatrix} ah_{1,1} & ah_{2,1} & bh_{1,1} & bh_{2,1} \\ ah_{1,2} & ah_{2,2} & bh_{1,2} & bh_{2,2} \\ \vdots & \vdots & \vdots & \vdots \\ ah_{1,n_R} & ah_{2,n_R} & bh_{1,n_R} & bh_{2,n_R} \\ c^*h^*_{2,1} & -c^*h^*_{1,1} & d^*h^*_{2,1} & -d^*h^*_{1,1} \\ c^*h^*_{2,2} & -c^*h^*_{1,2} & d^*h^*_{2,2} & -d^*h^*_{1,2} \\ \vdots & \vdots & \vdots & \vdots \\ c^*h^*_{2,n_R} & -c^*h^*_{1,n_R} & d^*h^*_{2,n_R} & -d^*h^*_{1,n_R} \end{pmatrix}, \quad \mathbb{H}_2 = \begin{pmatrix} ah_{3,1} & ah_{4,1} & bh_{3,1} & bh_{4,1} \\ ah_{3,2} & ah_{4,2} & bh_{3,2} & bh_{4,2} \\ \vdots & \vdots & \vdots & \vdots \\ ah_{3,n_R} & ah_{4,n_R} & bh_{3,n_R} & bh_{4,n_R} \\ c^*h^*_{4,1} & -c^*h^*_{3,1} & d^*h^*_{4,1} & -d^*h^*_{3,1} \\ c^*h^*_{4,2} & -c^*h^*_{3,2} & d^*h^*_{4,2} & -d^*h^*_{3,2} \\ \vdots & \vdots & \vdots & \vdots \\ c^*h^*_{4,n_R} & -c^*h^*_{3,n_R} & d^*h^*_{4,n_R} & -d^*h^*_{3,n_R} \end{pmatrix},$$

$$\mathbb{H}_3 = \begin{pmatrix} ah_{2,1}\varphi & ah_{3,1}\varphi & bh_{2,1}\varphi & bh_{3,1}\varphi \\ ah_{2,2}\varphi & ah_{3,2}\varphi & bh_{2,2}\varphi & bh_{3,2}\varphi \\ \vdots & \vdots & \vdots & \vdots \\ ah_{2,n_R}\varphi & ah_{3,n_R}\varphi & bh_{2,n_R}\varphi & bh_{3,n_R}\varphi \\ c^*h^*_{3,1}\varphi^* & -c^*h^*_{2,1}\varphi^* & d^*h^*_{3,1}\varphi^* & -d^*h^*_{2,1}\varphi^* \\ c^*h^*_{3,2}\varphi^* & -c^*h^*_{2,2}\varphi^* & d^*h^*_{3,2}\varphi^* & -d^*h^*_{2,2}\varphi^* \\ \vdots & \vdots & \vdots & \vdots \\ c^*h^*_{3,n_R}\varphi^* & -c^*h^*_{2,n_R}\varphi^* & d^*h^*_{3,n_R}\varphi^* & -d^*h^*_{2,n_R}\varphi^* \end{pmatrix},$$

$$\mathbb{H}_4 = \begin{pmatrix} ah_{4,1}\varphi & ah_{1,1}\varphi & bh_{4,1}\varphi & bh_{1,1}\varphi \\ ah_{4,2}\varphi & ah_{1,2}\varphi & bh_{4,2}\varphi & bh_{1,2}\varphi \\ \vdots & \vdots & \vdots & \vdots \\ ah_{4,n_R}\varphi & ah_{1,n_R}\varphi & bh_{4,n_R}\varphi & bh_{1,n_R}\varphi \\ c^*h^*_{1,1}\varphi^* & -c^*h^*_{4,1}\varphi^* & d^*h^*_{1,1}\varphi^* & -d^*h^*_{4,1}\varphi^* \\ c^*h^*_{1,2}\varphi^* & -c^*h^*_{4,2}\varphi^* & d^*h^*_{1,2}\varphi^* & -d^*h^*_{4,2}\varphi^* \\ \vdots & \vdots & \vdots & \vdots \\ c^*h^*_{1,n_R}\varphi^* & -c^*h^*_{4,n_R}\varphi^* & d^*h^*_{1,n_R}\varphi^* & -d^*h^*_{4,n_R}\varphi^* \end{pmatrix},$$

where $h_{i,j}$ is the channel fading coefficient between the ith transmit antenna and the jth receive antenna, and $\varphi = \mathrm{e}^{j\phi}$.

Let $\mathbf{h}_{1,\ell}$, $\mathbf{h}_{2,\ell}$, $\mathbf{h}_{3,\ell}$ and $\mathbf{h}_{4,\ell}$ denote the columns of $\mathbb{H}_\ell$. Since the high-rate STBC is constructed from a linear combination of two Alamouti codes, the orthogonal property exists for two pairs of the columns of $\mathbb{H}_\ell$, namely $\langle \mathbf{h}_{1,\ell}, \mathbf{h}_{2,\ell} \rangle = \langle \mathbf{h}_{3,\ell}, \mathbf{h}_{4,\ell} \rangle = 0$. Based on this

property, the ML detection can be simplified. Specifically, for a specific $\mathbb{H}_\ell$, the ML detection in (3.4) can be rewritten as

$$(\hat{x}_{1,\ell}, \hat{x}_{2,\ell}, \hat{x}_{3,\ell}, \hat{x}_{4,\ell}) = \underset{x_i \in \Psi}{\operatorname{argmin}} \left\| \mathbf{y} - \sqrt{\frac{\rho}{\mu}} \mathbb{H}_\ell \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \right\|^2. \tag{3.8}$$

Because the column vectors $\mathbf{h}_{1,\ell}$ and $\mathbf{h}_{2,\ell}$ are orthogonal, for given values of (x_3, x_4), the ML estimates of x_1 and x_2 can be performed independently as follows:

$$(\check{x}_{1,\ell}|_{x_3,x_4}) = \underset{x_1 \in \Psi}{\operatorname{argmin}} \left\| \mathbf{y} - \sqrt{\frac{\rho}{\mu}} (\mathbf{h}_{1,\ell} x_1 + \mathbf{h}_{3,\ell} x_3 + \mathbf{h}_{4,\ell} x_4) \right\|^2 \tag{3.9}$$

$$(\check{x}_{2,\ell}|_{x_3,x_4}) = \underset{x_2 \in \Psi}{\operatorname{argmin}} \left\| \mathbf{y} - \sqrt{\frac{\rho}{\mu}} (\mathbf{h}_{2,\ell} x_2 + \mathbf{h}_{3,\ell} x_3 + \mathbf{h}_{4,\ell} x_4) \right\|^2. \tag{3.10}$$

After collecting the results from (3.9) and (3.10), which are expressed as $(\check{x}_{1,\ell}, \check{x}_{2,\ell}|_{x_3,x_4})$, the ML estimates $(\hat{x}_{1,\ell}, \hat{x}_{2,\ell}, \hat{x}_{3,\ell}, \hat{x}_{4,\ell})$ will be then determined by (3.8) over all $((\check{x}_{1,\ell}, \check{x}_{2,\ell}|_{x_3,x_4}), x_3, x_4)$ values[2]. Since the above ML estimations are performed for a particular $\mathbb{H}_\ell$, the receiver makes a final decision by choosing the minimum antenna combination metric $\hat{\ell} = \arg \min_{\ell} m_\ell$, $\ell = 1, 2, \cdots, q$, where m_ℓ is the value of the minimum metric in (3.8).

Compared to the ML decoding of the STBC-SM scheme that has a complexity of $2qM$, the decoding complexity[3] of the HR-STBC-SM scheme is $2qM^2$, which is higher for the same value of M. Fortunately, for the same spectral efficiency in terms of bits/s/Hz the HR-STBC-SM uses a lower-order constellation and it turns out that the ML detection complexity of the HR-STBC-SM scheme could be comparable to that of the STBC-SM scheme.

Before closing this section, an upper bound on the error probability is given as it shall be used to gauge the performance obtained by computer simulation. First, the pairwise error

[2]Alternatively, because of the orthogonal property of $\mathbf{h}_{3,\ell}$ and $\mathbf{h}_{4,\ell}$, the ML detection can be performed the other way round with the same complexity where the receiver first estimates x_3 and x_4 independently for each known pair (x_1, x_2).

[3]Note that the detections of $\hat{x}_{1,\ell}$, and $\hat{x}_{2,\ell}$ in (3.9) and (3.10) only require simple threshold circuits for given (x_3, x_4).

probability for deciding on codeword $\mathbf{X}_j$ given that $\mathbf{X}_i$ was transmitted is given by [36]

$$\Pr\left(\mathbf{X}_i \to \mathbf{X}_j\right) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \left(\frac{1}{1+\frac{\rho\lambda_1}{4\sin^2\vartheta}}\right)^{n_R} \left(\frac{1}{1+\frac{\rho\lambda_2}{4\sin^2\vartheta}}\right)^{n_R} d\vartheta \tag{3.11}$$

where λ_1 and λ_2 are the eigenvalues of matrix $(\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_j)^H$ under the normalization $\mu = 1$ and $E\{\mathrm{tr}(\mathbf{X}^H\mathbf{X})\} = 2$. Assume that k bits are transmitted over two consecutive symbol intervals, the union bound on the bit error probability is

$$P(\text{error}) \leq \frac{1}{2^k}\sum_{i=1}^{2^k}\sum_{j=1}^{2^k} \frac{\Pr\left(\mathbf{X}_i \to \mathbf{X}_j\right)\chi\left(\mathbf{X}_i, \mathbf{X}_j\right)}{k} \tag{3.12}$$

where $\chi\left(\mathbf{X}_i, \mathbf{X}_j\right)$ is the number of bits in error when comparing matrices $\mathbf{X}_i$ and $\mathbf{X}_j$.

3.1.3 Simulation Results and Comparison

In this section, the BER simulation results of the HR-STBC-SM and STBC-SM schemes are presented and compared for various numbers of transmit antennas and spectral efficiency values versus the average SNR per receive antenna (ρ). In all simulations, four receive antennas are employed.

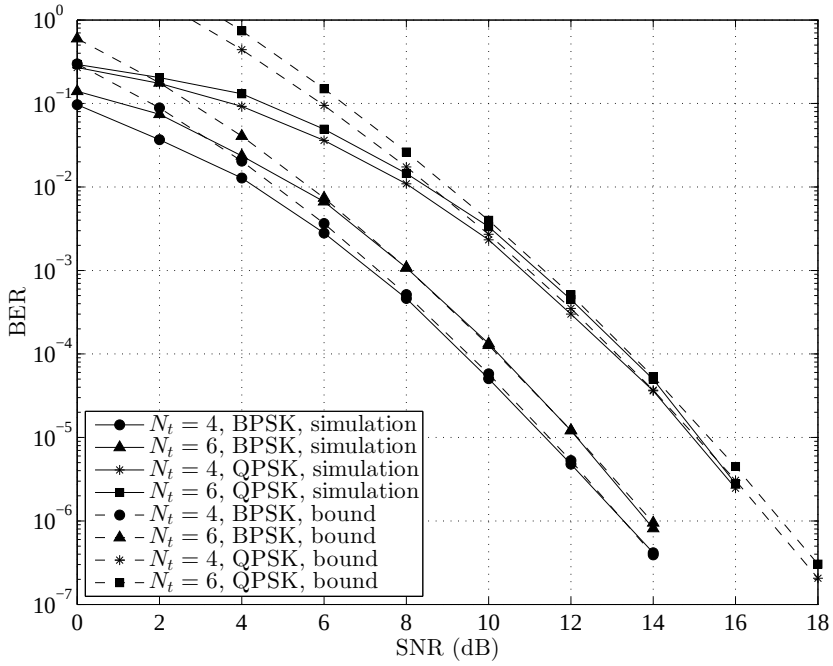

Figure 3.2 Performance comparison between theoretical upper bound and simulation results of the HR-STBC-SM scheme.

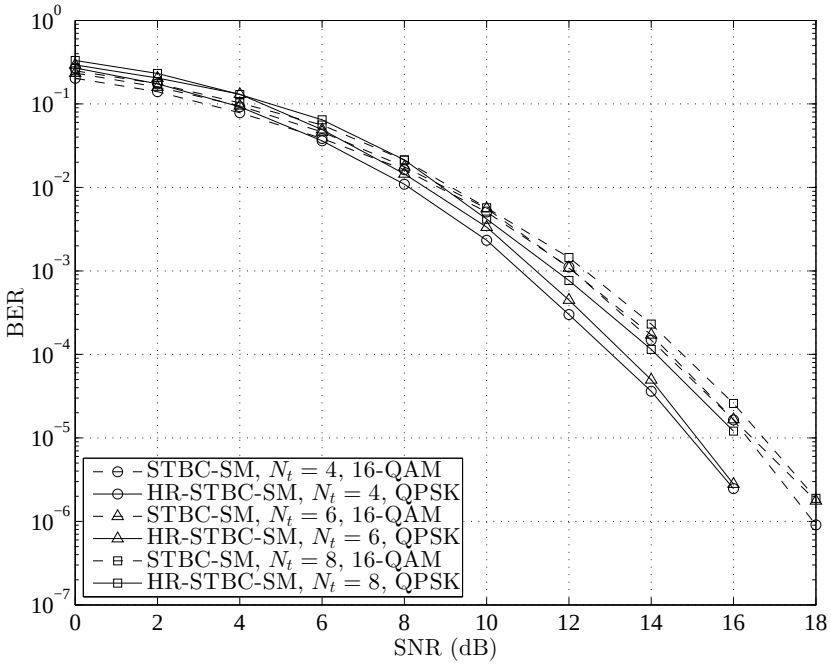

Figure 3.3 BER comparison between HR-STBC-SM and STBC-SM schemes at 5 bits/s/Hz, 5.5 bits/s/Hz and 6 bits/s/Hz.

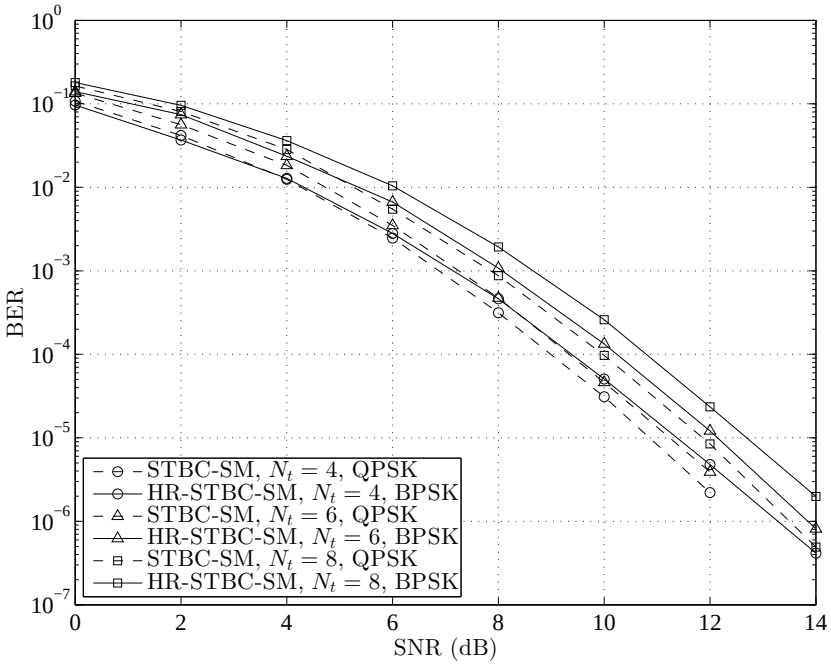

Figure 3.4 BER comparison between HR-STBC-SM and STBC-SM schemes at 3 bits/s/Hz, 3.5 bits/s/Hz and 4 bits/s/Hz.

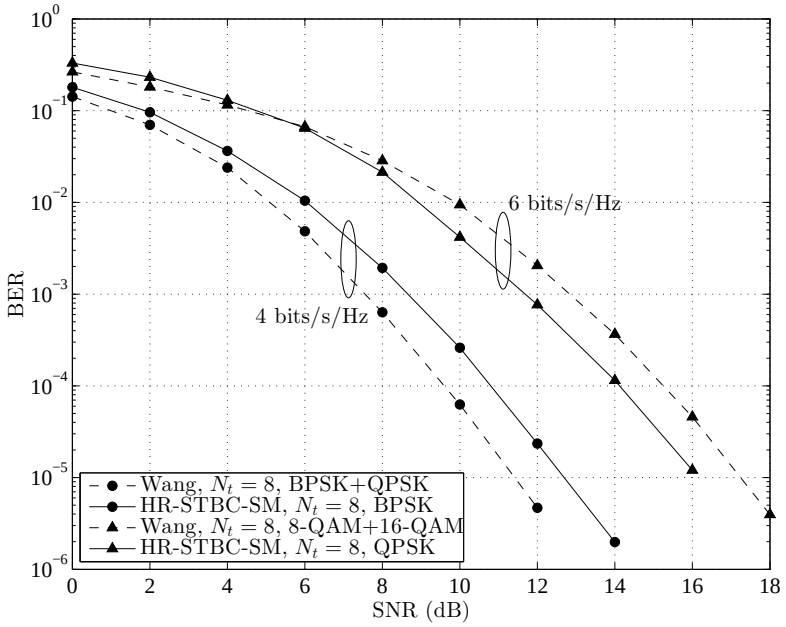

Figure 3.5 BER comparison between HR-STBC-SM and Wang's schemes at 4 bits/s/Hz and 6 bits/s/Hz.

First, Figure 3.2 compares the upper bound in (3.12) and the BER performance obtained by simulation for the cases of 4 and 6 antennas with BPSK and QPSK constellations. The figure clearly illustrates the tightness of the union bound at high SNR, which makes it useful to study the error performance behavior of the proposed HR-STBC-SM scheme with different system setups.

Figure 3.3 shows the BER performance comparison between the HR-STBC-SM and STBC-SM schemes at spectral efficiencies of 5, 5.5 and 6 bits/s/Hz, which correspond to systems with 4, 6 and 8 antennas. To deliver such spectral efficiencies, QPSK is used for the HR-STBC-SM scheme, whereas 16-QAM is used for the STBC-SM scheme. More importantly, at 5 bits/s/Hz and BER level of 10^{-5}, the HR-STBC-SM scheme provides a 1.8dB SNR gain over the STBC-SM scheme. Similarly, for the cases of 6 and 8 antennas (5.5 and 6 bits/s/Hz), the SNR gains are 1.6dB and 0.8dB, respectively. Such SNR gains are predicted by the analysis and comparison of coding gains in Table 3.1.

In Figure 3.4, the BER curves of STBC-SM and HR-STBC-SM with 4, 6 and 8 antennas schemes are compared at lower spectral efficiencies. Specifically, QPSK is used for STBC-

SM while BPSK is used for HR-STBC-SM. The corresponding spectral efficiencies are 3, 3.5 and 4 bits/s/Hz, respectively. It can be seen that the HR-STBC-SM scheme performs quite similar to the STBC-SM scheme at 3 bits/s/Hz. At the spectral efficiency of 4 bits/s/Hz (with 8 available antennas), the STBC-SM actually outperforms our proposed scheme by 1 dB. Again, this can be predicted from Table 3.1, which shows that the coding gain of the STBC-SM scheme is 1.2179, while that of the HR-STBC-SM scheme is 0.5858.

In Figure 3.5, the BER curves of the HR-STBC-SM and the scheme proposed by Wang *et al.* in [37] are evaluated at 4 and 6 bits/s/Hz. Wang's scheme uses a $(4, 3)$ error-correcting code together with 8 transmit antennas to create 32 codewords, which is larger than 16 codewords of HR-STBC-SM. As can be seen from the figure, at 4 bits/s/Hz, Wang's scheme has about 1dB SNR gain as compared to the proposed HR-STBC-SM. However, at 6 bits/s/Hz, the HR-STBC-SM scheme achieves a 1 dB gain over Wang's scheme.

3.2 Bi-Generalised Space Shift Keying over MIMO Channels

To further reduce the complexity of decoding in SM, amplitude/phase modulation symbols are removed and the information source is solely conveyed by antenna indices. This method is referred to as space shift keying (SSK) [32]. Although having very low decoding complexity, the drawback of SSK is low spectral efficiency. There are many recent studies on increasing the spectral efficiency of SSK. For example, the generalized SSK (GSSK) [40] activates a fix number of antennas that is larger than one to yield a larger number of codewords. The spectral efficiency of GSSK is $\left\lfloor \log_2 \binom{N_t}{n_t} \right\rfloor$ bits/sec/Hz, where N_t is the number of available transmit antennas while n_t is the number of activated antennas. Bi-space shift keying (BiSSK) [41] is another scheme that doubles the rate of SSK by multiplexing two orthogonal channels: an in-phase channel (corresponding to real numbers in baseband) and a quadrature channel (corresponding to imaginary numbers). The transmission on each channel is determined by the SSK rule. Recently, another scheme that also uses the multiplexing technique to enhance the spectral efficiency is layered-SSK (LSSK) [42]. In LSSK, the layers are determined by the SSK rule and multiplexing is performed by adding/dropping active antennas.

The new SSK scheme proposed in this research[4] is also based on the multiplexing technique. Compared to BiSSK, the new scheme, termed bi-generalized SSK (BiGSSK), employs GSSK rule for each inphase/quadrature data stream. As such, BiGSSK has the same advantage of low detection complexity as with SSK, while doubling the transmission rate of GSSK. By optimizing the signals transmitted over activated antennas, it is shown that such spectral efficiency advantage of BiGSSK is compromised only by a negligible performance loss. Numerical results and performance comparison with GSSK, BiSSK, LSSK demonstrate the efficiency of the proposed BiGSSK.

3.2.1 System Model and Performance Analysis

Consider a MIMO system employing SSK-based modulation that is equipped with N_t transmit and N_r receive antennas. The baseband input/output signal model is given by

$$\mathbf{y} = \sqrt{E_s}\mathbf{H}\mathbf{x} + \mathbf{n}. \tag{3.13}$$

In the above expression, $\mathbf{x}$ is a $N_t \times 1$ transmit symbol vector comprising of nonzero elements (corresponding to activated antennas) and zero elements (corresponding to idle antennas). The transmit symbol vector $\mathbf{x}$ is drawn with equal probabilities from a codebook A and $\mathrm{E}\{\|\mathbf{x}\|^2\} = 1$. The $N_r \times 1$ vector $\mathbf{y}$ is the receive signal vector, while $\mathbf{n}$ is $N_r \times 1$ noise vector whose entries are independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and variance N_0. The $N_r \times N_t$ matrix $\mathbf{H}$ is the channel matrix, whose entries are i.i.d complex Gaussian random variables with zero mean and unit variance. This means that the channel is flat Rayleigh fading. The constant E_s is the average transmitted energy per each symbol vector.

Given the signal model in (3.13), the ML detection of the transmitted vector $\mathbf{x}$ is to solve the following constrained least-square problem:

$$\tilde{\mathbf{x}}_{\mathrm{ML}} = \arg\min_{\mathbf{x}\in A} \left\| \mathbf{y} - \sqrt{E_s}\mathbf{H}\mathbf{x} \right\|^2 \tag{3.14}$$

[4]The contributions in this section are presented in [43].

The bit error probability can be approximated as

$$\mathrm{P}[\text{bit error}] \leq \frac{1}{\kappa} \sum_{i=1}^{\kappa} \sum_{\substack{j=1 \\ j \neq i}}^{\kappa} N_{i,j} \mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j), \tag{3.15}$$

where κ is the total number of codewords, $\mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j)$ is the pairwise error probability (PEP) of deciding on $\mathbf{x}_j$ given that $\mathbf{x}_i$ was transmitted, and $N_{i,j}$ is the number of bits in error when choosing $\mathbf{x}_j$ over $\mathbf{x}_i$. Conditioned on $\mathbf{H}$, the PEP is

$$\begin{aligned}
&\mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j \mid \mathbf{H}) \\
&= \mathrm{P}\left(\left\| \mathbf{y} - \sqrt{E_s}\mathbf{H}\mathbf{x}_i \right\|^2 > \left\| \mathbf{y} - \sqrt{E_s}\mathbf{H}\mathbf{x}_j \right\|^2 \middle| \mathbf{H} \right) \\
&= \mathrm{P}\left(\Re(\mathbf{n}^H(\mathbf{H}\mathbf{x}_i - \mathbf{H}\mathbf{x}_j)) > \frac{\sqrt{E_s}}{2} \left\| \mathbf{H}\mathbf{x}_i - \mathbf{H}\mathbf{x}_j \right\|^2 \middle| \mathbf{H} \right) \\
&= Q\left(\sqrt{\frac{E_s}{2N_0} \left\| \mathbf{H}\mathbf{x}_i - \mathbf{H}\mathbf{x}_j \right\|^2} \right) = Q\left(\sqrt{\sum_{n=1}^{2N_r} \epsilon_n^2} \right),
\end{aligned} \tag{3.16}$$

where ϵ_n is a Gaussian random variable with zero mean and variance $\sigma_{ij}^2 = \frac{E_s}{4N_0} \left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2$. It follows that the random variable $\zeta_{ij} = \sum_{n=1}^{N_r} \epsilon_n^2$ obeys a Chi-squared distribution with $2N_r$ degrees of freedom, whose pdf is given by [44]

$$f(\zeta_{ij}) = \frac{\zeta_{ij}^{N_r - 1} \exp(-\frac{\zeta_{ij}}{2\sigma_{ij}^2})}{(2\sigma_{ij}^2)^{N_r} \Gamma(N_r)}, \tag{3.17}$$

where $\Gamma(\cdot)$ is the Gamma function. The unconditioned PEP is hence obtained as

$$\mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j) = \mathrm{E}[\mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j | \mathbf{H})] = \int_0^\infty Q\left(\sqrt{\zeta_{ij}} \right) f(\zeta_{ij}) \, \mathrm{d}\zeta_{ij} \tag{3.18}$$

which has a closed-form expression as [45]:

$$\mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j) = \delta_{ij}^{N_r} \sum_{m=0}^{N_r} \binom{N_r - 1 + m}{m} [1 - \delta_{ij}]^m, \tag{3.19}$$

where $\delta_{ij} = \frac{1}{2}\left(1 - \sqrt{1 - \frac{1}{1+\sigma_{ij}^2}} \right)$. Substituting (3.19) into (3.15) give an upper bound of the bit error probability for a SSK-based modulation scheme.

To gain a better understanding of the error performance of a SSK-based modulation scheme, examine the following looser upper bound [46]:

$$\mathrm{P}(\mathbf{x}_i \to \mathbf{x}_j) \leq \frac{1}{2} \left(\sigma_{ij}^2 + 1 \right)^{-N_r} \leq a \left(\frac{E_s}{N_0} \right)^{-N_r} \left(\left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2 \right)^{-N_r} \tag{3.20}$$

where $a = 4^{N_r}/2$. The above expression clearly indicates that the error probability depends on the Euclidean distances among the possible transmit symbol vectors. In the special case that each transmit symbol vector contains only elements 1 and 0 (e.g. in SSK or GSSK), the Euclidean distances are the same as the Hamming distances.

3.2.2 Proposed Bi-Generalized Space Shift Keying

The analysis in the previous section shows that the performance of a SSK-based modulation scheme depends on the Euclidean distance between any two codewords (i.e., two transmit symbol vectors). As discussed before, the BiSSK scheme multiplexes one real number $(+1)$ and one imaginary number $(+j)$ based on two input bits where the indices of the active antennas carrying those real and imaginary numbers are determined by the SSK rule. Different from BiSSK, in order to improve the spectral efficiency, the proposed BiGSSK multiplexes multiple real and imaginary numbers based on a group of $\lambda > 2$ information bits. Specifically, $\lambda/2$ bits select a subset of antennas to carry real numbers, while the other $\lambda/2$ bits select a subset of antennas to carry imaginary numbers. The selection of antenna subsets for either real or imaginary numbers follows the GSSK rule.

To illustrate the proposed scheme, Table 3.3 describes a partial codebook for the BiGSSK scheme with $N_t = 5$ and $n_t = 2$. This means that there are $\binom{5}{2} = 10$ different antenna subsets, hence $\lambda/2 = 3$ bits can be carried by either the real or imaginary data stream in each time slot. Note that Table 3.3 only shows codewords corresponding to the first 3 bits equal to $\mathbf{s}_1 = 000$, while the GSSK rule to select a subset of $n_t = 2$ active antennas out of $N_t = 5$ available antennas is as in Table 3.2. It is pointed out that when the GSSK rule is used with $n_t = 2$, the maximum number of active antennas of the BiGSSK scheme is 4.

Since the performance of a SSK-based scheme depends on the minimum distance between any two codewords, an improvement shall be made to the codebook in Table 3.3 to increase its minimum Euclidean distance. To this end, observe that the nonzero elements in Table 3.3, namely $(1+j)/2$, $1/2$ and $j/2$ are resulted directly by multiplexing two inphase/quadrature GSSK data streams. However, from the perspective of diverting total codeword energy over transmit antennas, these are simply three different complex numbers showing how energies

Table 3.2 GSSK rule for selecting antenna subsets with $N_t = 5$, $n_t = 2$.

Information bits	GSSK symbol vector
000	$[1, 1, 0, 0, 0]^T$
001	$[1, 0, 1, 0, 0]^T$
010	$[1, 0, 0, 1, 0]^T$
011	$[1, 0, 0, 0, 1]^T$
100	$[0, 1, 1, 0, 0]^T$
101	$[0, 1, 0, 1, 0]^T$
110	$[0, 1, 0, 0, 1]^T$
111	$[0, 0, 1, 1, 0]^T$

are transmitted over different active antennas. Since the minimum distance of the codebook directly depends on the distances among the nonzero elements, the improvement is to replace $(1 + j)/2$, $1/2$ and $j/2$ by three numbers that have the same energy (i.e., magnitude) and maximally spaced in the complex plane, i.e., they lie on the vertices of an equilateral triangle. One set of such numbers are $-(\sqrt{3} + j)$, $\sqrt{3} - j$, and $2j$. Let n_{RF} be the *average* number of active antennas (i.e., the average number of RF chains), then it is simple to show that when the nonzero elements in the codebook have the same magnitude, the *normalized* minimum squared distance of the proposed codebook is $d^2_{\min} = \frac{2}{n_{\mathrm{RF}}}$.

Proposition 1. *The average number of active antennas in Table 3.3 is* $3 \leq n_{\mathrm{RF}} < 4$.

Proof. With the GSSK rule in Table 3.2, among codewords with the same information bits $\mathbf{s}_1$ but different $\mathbf{s}_2$, there is always only one codeword that activates two antennas and not less than one codeword that activate four antennas, while codewords with three activated antennas always exist. Hence, the average number of RF chains used is $3 \leq n_{\mathrm{RF}} < 4$. The above observation also holds for other cases of $\mathbf{s}_1$. □

It follows from the above Proposition that the minimum squared distance of the BiGSSK is bounded by $\frac{1}{2} < d^2_{\min} = \frac{2}{n_{\mathrm{RF}}} \leq \frac{2}{3}$. In general, $d^2_{\min}$ approaches the upper bound $\frac{2}{3}$ when N_t is small ($N_t \leq 8$) since the codewords with three activated antennas dominate those with

Table 3.3 Partial codebook in BiGSSK with the first three information bits being $\mathbf{s}_1 = 000$, $N_t = 5$.

Information bits $\mathbf{s}_1 = 000$, $\mathbf{s}_2$	The symbol vector $\mathbf{x}$
000,000	$[1+j, 1+j, 0, 0, 0]^T/2$
000,001	$[1+j, 1, j, 0, 0]^T/2$
000,010	$[1+j, 1, 0, j, 0]^T/2$
000,011	$[1+j, 1, 0, 0, j]^T/2$
000,100	$[1, 1+j, j, 0, 0]^T/2$
000,101	$[1, 1+j, 0, j, 0]^T/2$
000,110	$[1, 1+j, 0, 0, j]^T/2$
000,111	$[1, 1, j, j, 0]^T/2$

four activated antennas. On the other hand, $d^2_{\min}$ tends to $\frac{1}{2}$ as N_t becomes large. This reduction in $d^2_{\min}$ is compensated by an increase in the transmission rate.

In general, the proposed BiGSSK scheme is constructed as follows:

1. The input bits sequence is divided into two equal-sized groups, $\mathbf{s}_1$ and $\mathbf{s}_2$.
2. Using the GSSK mapping rule with $n_t = 2$, $\mathbf{s}_1$ and $\mathbf{s}_2$ form two $N_t \times 1$ symbol vectors $\mathbf{x}_R$ and $\mathbf{x}_I$.
3. Sum $\mathbf{x}_R$ and $j\mathbf{x}_I$ and normalize the sum to form an intermediate transmit symbol vector $\bar{\mathbf{x}} = (\mathbf{x}_R + j\mathbf{x}_I)/2$.
4. Change the values $(1+j)/2$, $1/2$ and $j/2$ in $\bar{\mathbf{x}}$ to $-\frac{\sqrt{3}+j}{\sqrt{\alpha}}$, $\frac{\sqrt{3}-j}{\sqrt{\alpha}}$, and $\frac{2j}{\sqrt{\alpha}}$ to obtain the final transmit symbol vector $\mathbf{x}$, where $\sqrt{\alpha} = \sqrt{4n_{\mathrm{RF}}}$ is the normalizing factor to ensure that $\mathrm{E}\{\|\mathbf{x}\|^2\} = 1$.

As an example, consider a MIMO system with $N_t = 4$. Suppose that a group of four information bits to be transmitted is 1001. First, determine $\mathbf{s}_1 = [1, 0]$ and $\mathbf{s}_2 = [0, 1]$. Using the GSSK mapping rule with $n_t = 2$ as in Table 3.4, $\mathbf{s}_1$ and $\mathbf{s}_2$ form vectors $x_R = [1, 0, 0, 1]$ and $x_I = [0, 0, 1, 1]$. Summing $\mathbf{x}_R$ and $j\mathbf{x}_I$ and normalizing form vector $\bar{\mathbf{x}} = [1, 0, j, 1+j]/2$.

Changing elements $1/2$ to $\frac{\sqrt{3}-j}{2\sqrt{3}}$, $j/2$ to $\frac{2j}{2\sqrt{3}}$, and $(1+j)/2$ to $-\frac{\sqrt{3}+j}{2\sqrt{3}}$ (here $\alpha = 12$ and $n_{\mathrm{RF}} = 3$) gives the transmit symbol vector $\mathbf{x} = [\sqrt{3}-j, 0, 2j, -\sqrt{3}-j]/2\sqrt{3}$. The full symbol mapping for this example is summarized in Table 3.5.

Table 3.4 GSSK symbol mapping with $N_t = 4$, $n_t = 2$.

Information bits	GSSK symbol vector
00	$[1,1,0,0]^T$
01	$[0,0,1,1]^T$
10	$[1,0,0,1]^T$
11	$[0,1,1,0]^T$

It is pointed out that the GSSK rule used in the proposed BiGSSK can be applied with $n_t \geq 2$. However, the performance will be degraded with increasing n_t since $d^2_{\min} = \frac{2}{n_{\mathrm{RF}}}$ and n_{RF} obviously increases with n_t. The tradeoff between transmission rate, performance and system complexity in selecting n_t is a topic of further studies. This present research focuses on the GSSK with $n_t = 2$ as this selection of n_t minimizes the number of RF chains. In general, the spectral efficiency of BiGSSK is $2\left\lfloor \log_2\binom{N_t}{n_t} \right\rfloor$ bits/sec/Hz. Table 3.6 compares the minimum squared Euclidean distances along with the transmission rates between BiGSSK and other SSK-based modulation schemes.

3.2.3 Numerical Results

In this section, the BER simulation results of the BiGSSK and SSK schemes are presented and compared for various numbers of transmit antennas and spectral efficiency values versus the average SNR per bit per receive antenna (E_b/N_0). In all simulations, three receive antennas are employed.

As can be seen from (3.20), the performance of a SSK-based modulation scheme is dominated by the worst-case error, which is determined by parameter $\frac{E_s}{N_0} d^2_{\min} = \frac{\lambda E_b}{N_0} d^2_{\min}$, where λ is the number of bits carried by each symbol vector. The BERs of BiSSK and BiGSSK with different system settings are shown in Figure 3.6. The value $d^2_{\min}$ of BiSSK is 1, which is larger than $d^2_{\min}$ of BiGSSK. However the number of bits transmitted per codeword in

Table 3.5 BiGSSK symbol mapping with $N_t = 4$.

Information bits $\mathbf{s}_1$, $\mathbf{s}_2$	BiGSSK symbol vector $\mathbf{x}$
00,00	$[-\sqrt{3}-j, -\sqrt{3}-j, 0, 0]^T/2\sqrt{3}$
00,01	$[\sqrt{3}-j, \sqrt{3}-j, 2j, 2j]^T/2\sqrt{3}$
00,10	$[-\sqrt{3}-j, \sqrt{3}-j, 0, 2j]^T/2\sqrt{3}$
00,11	$[\sqrt{3}-j, -\sqrt{3}-j, 2j, 0]^T/2\sqrt{3}$
01,00	$[2j, 2j, \sqrt{3}-j, \sqrt{3}-j]^T/2\sqrt{3}$
01,01	$[0, 0, -\sqrt{3}-j, -\sqrt{3}-j]^T/2\sqrt{3}$
01,10	$[2j, 0, \sqrt{3}-j, -\sqrt{3}-j]^T/2\sqrt{3}$
01,11	$[0, 2j, -\sqrt{3}-j, \sqrt{3}-j]^T/2\sqrt{3}$
10,00	$[-\sqrt{3}-j, 2j, 0, \sqrt{3}-j]^T/2\sqrt{3}$
10,01	$[\sqrt{3}-j, 0, 2j, -\sqrt{3}-j]^T/2\sqrt{3}$
10,10	$[-\sqrt{3}-j, 0, 0, -\sqrt{3}-j]^T/2\sqrt{3}$
10,11	$[\sqrt{3}-j, 2j, 2j, \sqrt{3}-j]^T/2\sqrt{3}$
11,00	$[2j, -\sqrt{3}-j, \sqrt{3}-j, 0]^T/2\sqrt{3}$
11,01	$[0, \sqrt{3}-j, -\sqrt{3}-j, 2j]^T/2\sqrt{3}$
11,10	$[2j, \sqrt{3}-j, \sqrt{3}-j, 2j]^T/2\sqrt{3}$
11,11	$[0, -\sqrt{3}-j, -\sqrt{3}-j, 0]^T/2\sqrt{3}$

BiGSSK is higher than that with BiSSK for the same number of transmit antennas. This means that the PEPs of BiSSK and BiGSSK could still be comparable. At the fixed target transmission rates of 6, 8 and 10 bits/sec/Hz, BiGSSK shows a negligible performance loss (0.5dB) but using much smaller numbers of transmit antennas as compared to BiSSK (e.g. for 10 bits/sec/Hz, BiGSSK uses 9 antennas while BiSSK needs 32 antennas). With similar numbers of transmit antennas, the transmission rate of BiGSSK(7,3) is higher, by 2 bits/sec/Hz (which is 33.33% higher), than that of BiSSK(8,3) at the expense of 1dB performance loss at the BER level of 10^{-5}.

In Figure 3.7 the BER performances of BiGSSK and GSSK are compared for various settings. At a fixed number of transmit antennas, e.g. for 5 and 7 transmit antennas,

Table 3.6 Comparison in rate (bits/sec/Hz) and minimum squared distance between BiGSSK and other SSK-based modulation schemes.

N_t	4		5		6		7		16		32	
	Rate	$d^2_{\min}$	Rate	$d^2_{\min}$	Rate	$d^2_{\min}$	Rate	$d^2_{\min}$	Rate	$d^2_{\min}$	Rate	$d^2_{\min}$
SSK	2	2	2	2	2	2	2	2	4	2	5	2
GSSK, $n_t = 2$	2	1	3	1	3	1	4	1	6	1	8	1
BiSSK	4	1	4	1	4	1	4	1	8	1	10	1
BiGSSK	4	$\frac{2}{3}$	6	$\frac{2}{3.16}$	6	$\frac{2}{3.16}$	8	$\frac{2}{3.38}$	12	$\frac{2}{3.67}$	16	$\frac{2}{3.83}$

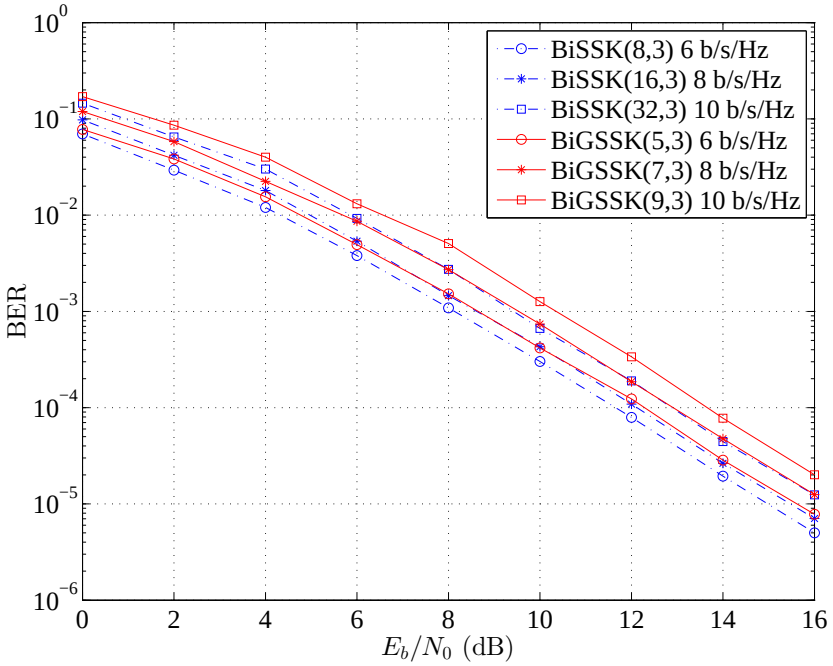

Figure 3.6 BER comparison of BiGSSK versus BiSSK with different (N_t, N_r).

BiGSSK doubles the transmission rate of GSSK at a negligible performance loss. On the other hand, for a fixed transmission rate, e.g. for 6 and 8 bits/sec/Hz, BiGSSK employs much smaller numbers of antennas while its BER performance is still comparable.

The simulation results for the same number of transmit antennas of LSSK(5,3) and BiGSSK(5,3) are shown in Figure 3.8. Both LSSK and BiGSSK schemes aim to enhance the transmission rate of SSK under a limited number of transmit antennas. As can be seen from the figure, the two schemes perform similarly at high E_b/N_0, but our proposed BiGSSK achieves a higher transmission rate (more than 2 bits/sec/Hz, or 50% higher) than LSSK.

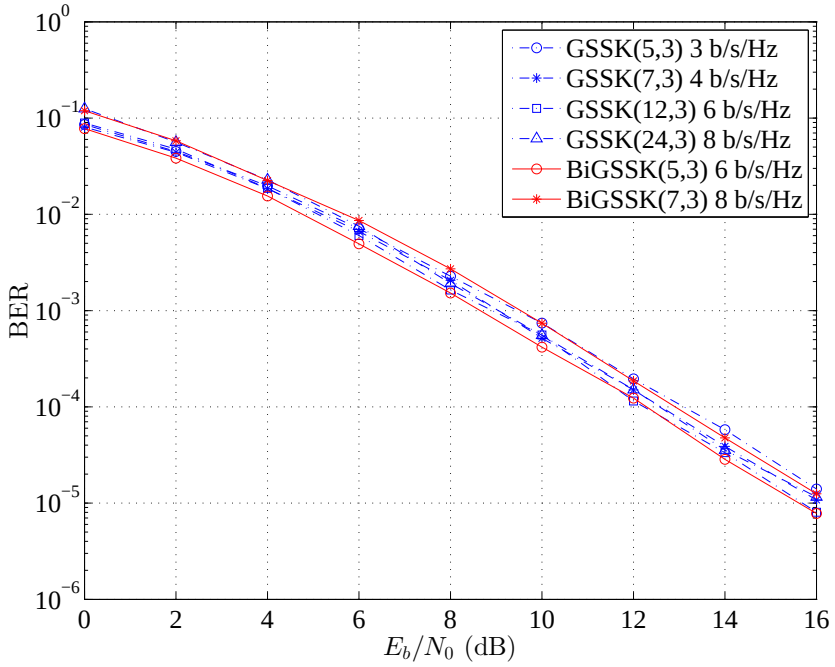

Figure 3.7 BER comparison of BiGSSK versus GSSK with different (N_t, N_r).

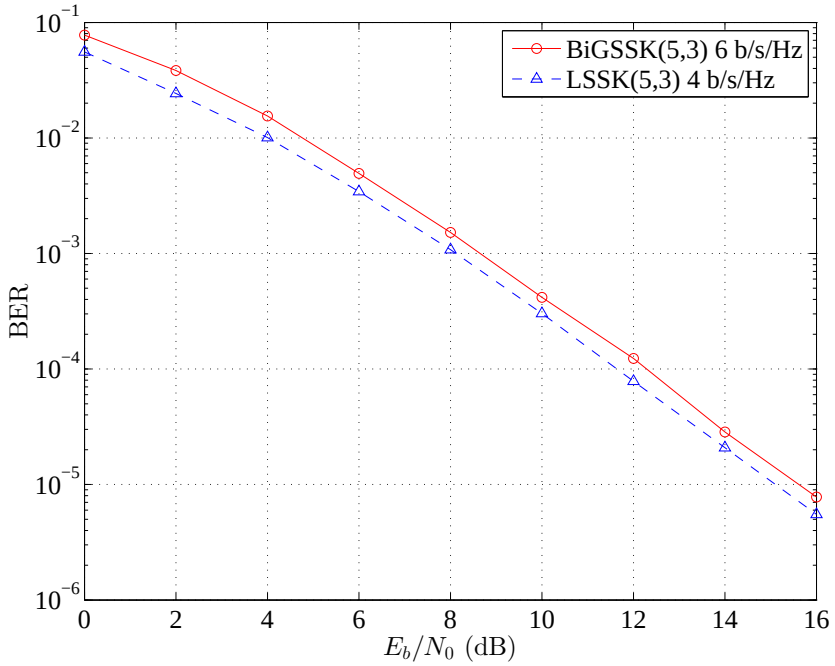

Figure 3.8 BER comparison of BiGSSK(N_t, N_r) and LSSK(N_t, N_r).

The similar performance of the two schemes can be predicted from the fact that $d^2_{\min}$ of LSSK(5,3) is 1, while $d^2_{\min}$ of BiGSSK(5,3) is $\frac{2}{3.1563}$.

3.3 Improved Quadrature Spatial Modulation

A major disadvantage of the conventional SM is that the data rate enhancement is only proportional to the base-two logarithm of the number of transmit antennas. This is unlike other spatial multiplexing techniques, such as V-BLAST, where the data rate increases linearly with the number of transmit antennas. Accordingly, research on enhancing the spectral efficiency of SM techniques has gained a significant interest in the literature.

In generalized SM (GSM) [47], a group of transmit antennas is considered a spatial constellation point and activated simultaneously to increase the overall spectral efficiency. A slight performance degradation of this scheme when compared to the conventional SM is reported in [47]. The conventional SM also has a drawback that it does not provide transmit diversity gain. In [36], space-time block coded spatial modulation (STBC-SM) is presented which achieves a transmit diversity order of two by activating two transmit antennas during two symbols duration and using the Alamouti code. In Section 3.1 the Alamouti code is replaced by a high-rate space-time block code, which improves coding gains at high data rates. Recently, quadrature spatial modulation (QSM) is proposed in [48] where the amplitude/phase modulation symbols are filtered and the real and imaginary parts are then transmitted by two different layers of antennas. By doing so, the number of bits conveyed by antenna index is doubled as compared to conventional SM. Another improvement proposed for SM, called enhanced SM (ESM) [49], uses a primary signal constellation when a single antenna is activated and other secondary constellations when two transmit antennas are activated. The flexibility in selecting one or two active antennas increases the number of bits conveyed by antenna index.

In this section, a novel SM scheme, which is called improved quadrature spatial modulation (IQSM), is investigated to further enhance the spectral efficiency of QSM. Different from STBC-SM where two symbols are transmitted during two symbol durations, the proposed scheme transmits two symbols in one symbol duration. Specifically, in IQSM the real and imaginary parts of two constellation symbols are transmitted on two different layers of antennas. Since two transmit antennas are activated for each in-phase or quadrature dimension, a larger number of codewords is generated, which means a higher spectral efficiency in terms

of bits/s/Hz. More importantly, to deliver a given spectral efficiency, it is possible to employ a lower-order constellation in IQSM, hence improving the system's error performance. To evaluate the performance of the proposed scheme and compare it with other existing schemes, an upper bound on the bit error probability under the maximum likelihood (ML) detection is obtained.

3.3.1 System Model and the Proposed Improved Quadrature Spatial Modulation

Consider a MIMO system that is equipped with N_t transmit and N_r receive antennas. The equivalent baseband input/output signal model is given by

$$\mathbf{y} = \sqrt{E_s}\mathbf{H}\mathbf{x} + \mathbf{n}. \tag{3.21}$$

In the above expression, $\mathbf{x}$ is an $N_t \times 1$ transmit symbol vector comprising of nonzero elements (corresponding to activated antennas) and zero elements (corresponding to idle antennas), $\mathbf{y}$ is an $N_r \times 1$ received signal vector, while $\mathbf{n}$ is an $N_r \times 1$ noise vector whose entries are independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and variance N_0. The $N_r \times N_t$ matrix $\mathbf{H}$ is the channel matrix, whose entries are i.i.d complex Gaussian random variables with zero mean and unit variance. This means that the channel is flat Rayleigh fading. The element on the mth row and nth column of $\mathbf{H}$ is denoted as $h_{m,n}$, which represents the channel path gain between the nth transmit and mth receive antennas. Each transmitted vector $\mathbf{x}$ has a unit norm and the constant E_s is interpreted as the average transmitted energy per each transmit symbol vector $\mathbf{x}$.

In QSM, the real and imaginary components of a constellation symbol are transmitted by two different layers of antennas (in-phase and quadrature dimensions). Each component (real or imaginary) has N_t possibilities to choose from. Therefore, a total of N_t^2 combinations (or codewords) are generated. With IQSM, the number of combinations for each dimension increases to $\binom{N_t}{2}$. This is possible because for each in-phase (or quadrature) dimension, two real (or imaginary) components of two symbols are transmitted on two different activated antennas. Thus, the total number of combinations for both dimensions is $\left[\binom{N_t}{2}\right]^2$.

The transmission framework of IQSM is as follows:

1. The input bit sequence is divided into two parts: antenna index part and constellation part. Bits in the antenna index part are further divided into 2 sub-parts: in-phase antenna index and quadrature antenna index.

2. Bits in the constellation part pick two constellation symbols a and b to be transmitted.

3. Bits in the in-phase antenna index sub-part determine 2 active antennas to transmit 2 real components $\mathrm{Re}(a)$ and $\mathrm{Re}(b)$. Let $\ell_{\mathrm{Re}(a)}$ and $\ell_{\mathrm{Re}(b)}$ denote indices of the 2 active antennas and without loss of generality, assume $\ell_{\mathrm{Re}(a)} < \ell_{\mathrm{Re}(b)}$. Thus the bits in the in-phase antenna index sub-part form a $N_t \times 1$ vector $\mathbf{x}_R$ where $\mathbf{x}_R(\ell_{\mathrm{Re}(a)}) = \mathrm{Re}(a)$ and $\mathbf{x}_R(\ell_{\mathrm{Re}(b)}) = \mathrm{Re}(b)$.

4. Similarly, bits in the quadrature antenna index sub-part determine 2 active antennas to transmit 2 imaginary components $\mathrm{Im}(a)$ and $\mathrm{Im}(b)$. Let $\ell_{\mathrm{Im}(a)}$ and $\ell_{\mathrm{Im}(b)}$ denote indices of the 2 active antennas and without loss of generality, assume $\ell_{\mathrm{Im}(a)} < \ell_{\mathrm{Im}(b)}$. The bits in the quadrature antenna index sub-part form a $N_t \times 1$ vector $\mathbf{x}_I$ where $\mathbf{x}_I(\ell_{\mathrm{Im}(a)}) = \mathrm{Im}(a)$ and $\mathbf{x}_I(\ell_{\mathrm{Im}(b)}) = \mathrm{Im}(b)$.

5. Finally, the transmit symbol vector is $\mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I$.

To illustrate the proposed IQSM, consider the transmission of eight information bits, 10010011, in a system with $N_t = 4$ and QPSK. First, these eight bits are divided into antenna index part $s_1 = [1, 0, 0, 1]$ and constellation part $s_2 = [0, 0, 1, 1]$. With QPSK and Gray mapping, the constellation part in s_2 selects 2 transmit symbols $a = 1+j$ and $b = -1-j$. On the other hand, the rule to activate 2 out of $N_t = 4$ antennas is given in Table 3.7. With such a rule, the first two bits $[1, 0]$ in s_1 select antennas Tx-1 and Tx-4 ($\ell_{\mathrm{Re}(a)} = 1$, $\ell_{\mathrm{Re}(b)} = 4$) for transmitting the real components of symbols a and b. Thus, $\mathbf{x}_R = [1, 0, 0, -1]^T$. Likewise, the remaining two bits $[0, 1]$ in s_1 select antennas Tx-1 and Tx-3 ($\ell_{\mathrm{Im}(a)} = 1$, $\ell_{\mathrm{Im}(b)} = 3$) for sending the quadrature components of a and b, hence $\mathbf{x}_I = [1, 0, -1, 0]^T$. The final transmit symbol vector is $\mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I = [1 + j, 0, -j, -1]^T$.

Table 3.7 Using 2 bits to select 2 out of $N_t = 4$ transmit antennas.

Indexing bits	Activated antennas
00	{1,2}
01	{1,3}
10	{1,4}
11	{2,3}

The spectral efficiency of the proposed IQSM is $\lambda = 2\left\lfloor \log_2 \binom{N_t}{2} \right\rfloor + 2\log_2 M$, where M is the constellation size ($M = 4$ for QPSK). It can be seen that the above spectral efficiency of IQSM is quite larger than $\frac{1}{2}\lfloor \log_2 \binom{N_t}{2} \rfloor + \log_2 M$ of STBC-SM and $2\lfloor \log_2 N_t \rfloor + \log_2 M$ of QSM or ESM. It is important to note that the number of RF chains used in IQSM is 4 while the number of RF chains in QSM and ESM is 2. This moderate increase of RF chains of the proposed IQSM is very reasonable, given the large improvement in the spectral efficiency when N_t is large. Furthermore, since the data rate of IQSM is now proportional to base-two logarithm of $\left[\binom{N_t}{2}\right]^2$, the scheme can be used in more general settings where the number of transmit antennas is not necessary a power of 2. This makes the proposed IQSM more flexible than the conventional SM and QSM.

3.3.2 Performance Analysis

Given the way the transmit symbol vector $\mathbf{x}$ is generated in the proposed IQSM as described in the previous section, the received signal vector can be re-written in the following form:

$$\begin{aligned}\mathbf{y} = \sqrt{E_s}[&\mathbf{h}_{\ell_{\mathrm{Re}(a)}}\mathrm{Re}(a) + \mathbf{h}_{\ell_{\mathrm{Re}(b)}}\mathrm{Re}(b) \\ &+ j\mathbf{h}_{\ell_{\mathrm{Im}(a)}}\mathrm{Im}(a) + j\mathbf{h}_{\ell_{\mathrm{Im}(b)}}\mathrm{Im}(b)] + \mathbf{n}. \end{aligned} \tag{3.22}$$

where $\ell_{\mathrm{Re}(a)},\ \ell_{\mathrm{Im}(a)},\ \ell_{\mathrm{Re}(b)},\ \ell_{\mathrm{Im}(b)} \in \{1, 2, \cdots, N_t\}$, $\ell_{\mathrm{Re}(a)} < \ell_{\mathrm{Re}(b)}$, $\ell_{\mathrm{Im}(a)} < \ell_{\mathrm{Im}(b)}$, and $\mathbf{h}_{\ell_{\mathrm{Re}(a)}}$ is the $\ell_{\mathrm{Re}(a)}$th column of $\mathbf{H}$. The maximum-likelihood (ML) detection of $\mathbf{x}$ is to solve:

$$\begin{aligned}
&[\tilde{\ell}_{\mathrm{Re}(a)}, \tilde{\ell}_{\mathrm{Im}(a)}, \tilde{\ell}_{\mathrm{Re}(b)}, \tilde{\ell}_{\mathrm{Im}(b)}, \tilde{\mathrm{Re}}(a), \tilde{\mathrm{Im}}(a), \tilde{\mathrm{Re}}(b), \tilde{\mathrm{Im}}(b)] \\
&= \arg \min_{\substack{\ell_{\mathrm{Re}(a)}, \ell_{\mathrm{Im}(a)}, \ell_{\mathrm{Re}(b)}, \ell_{\mathrm{Im}(b)} \\ \mathrm{Re}(a), \mathrm{Im}(a), \mathrm{Re}(b), \mathrm{Im}(b)}} \Big\| \mathbf{y} - \sqrt{E_s}[\mathbf{h}_{\ell_{\mathrm{Re}(a)}} \mathrm{Re}(a) + \\
&\qquad \mathbf{h}_{\ell_{\mathrm{Re}(b)}} \mathrm{Re}(b) + j\mathbf{h}_{\ell_{\mathrm{Im}(a)}} \mathrm{Im}(a) + j\mathbf{h}_{\ell_{\mathrm{Im}(b)}} \mathrm{Im}(b)] \Big\|^2 . \qquad (3.23)
\end{aligned}$$

There are 2^λ possible transmitted vectors. To evaluate the error performance of the ML detection, consider the pairwise error event that the transmitter sends vector $\mathbf{x}_m$, but the receiver decides on $\mathbf{x}_n$, $m \neq n$. Let a and b denote the two constellation symbols and $\ell_{\mathrm{Re}(a)},\ \ell_{\mathrm{Im}(a)},\ \ell_{\mathrm{Re}(b)},\ \ell_{\mathrm{Im}(b)}$ denote the antenna indices associated with vector $\mathbf{x}_m$. Likewise, let c and d denote the two constellation symbols and $\ell_{\mathrm{Re}(c)},\ \ell_{\mathrm{Im}(c)},\ \ell_{\mathrm{Re}(d)},\ \ell_{\mathrm{Im}(d)}$ denote the antenna indices associated with vector $\mathbf{x}_n$. Then the conditioned pairwire error probability (PEP) is given as

$$\Pr(\mathbf{x}_m \to \mathbf{x}_n \mid \mathbf{H}) = Q\left(\sqrt{\frac{E_s}{2N_0}} \|\mathbf{u}_m - \mathbf{u}_n\|\right), \qquad (3.24)$$

$$\begin{aligned}
\text{where} \quad \mathbf{u}_m &= [\mathbf{h}_{\ell_{\mathrm{Re}(a)}} \mathrm{Re}(a) + \mathbf{h}_{\ell_{\mathrm{Re}(b)}} \mathrm{Re}(b) \\
&+ j\mathbf{h}_{\ell_{\mathrm{Im}(a)}} \mathrm{Im}(a) + j\mathbf{h}_{\ell_{\mathrm{Im}(b)}} \mathrm{Im}(b)] \qquad (3.25) \\
\mathbf{u}_n &= [\mathbf{h}_{\ell_{\mathrm{Re}(c)}} \mathrm{Re}(c) + \mathbf{h}_{\ell_{\mathrm{Re}(d)}} \mathrm{Re}(d) \\
&+ j\mathbf{h}_{\ell_{\mathrm{Im}(c)}} \mathrm{Im}(c) + j\mathbf{h}_{\ell_{\mathrm{Im}(d)}} \mathrm{Im}(d)] \qquad (3.26)
\end{aligned}$$

Let $\gamma_k = |(\mathbf{u}_m - \mathbf{u}_n)_k|^2$, where $(\mathbf{u}_m - \mathbf{u}_n)_k$ is the kth element of vector $(\mathbf{u}_m - \mathbf{u}_n)$. Then it is not hard to see that, for a given pair $\{\mathbf{x}_m, \mathbf{x}_n\}$, γ_k is an exponential random variable with the same mean value $\bar{\gamma}_{m,n}$ for $1 \leq k \leq N_r$. By averaging $\Pr(\mathbf{x}_m \to \mathbf{x}_n \mid \mathbf{H})$ over the distribution of $\mathbf{H}$, the unconditioned PEP $\Pr(\mathbf{x}_m \to \mathbf{x}_n)$ is obtained. Instead of averaging over the distribution of $\mathbf{H}$, one can average the Q-function over the distributions of i.i.d.

exponential random variables γ_k. This gives the final PEP expression as follows:

$$\Pr(\mathbf{x}_m \to \mathbf{x}_n) = \mu_{m,n}^{N_r} \sum_{k=0}^{N_r-1} \binom{N_r - 1 + k}{k} [1 - \mu_{m,n}]^k \tag{3.27}$$

where $\mu_{m,n} = \left(1 - \sqrt{(\bar{\gamma}_{m,n}/2)/(1 + \bar{\gamma}_{m,n}/2)}\right)/2$. The average bit error probability (BEP) of the proposed IQSM can be evaluated using the following asymptotically tight union bound:

$$P(\text{error}) \leq \frac{1}{\lambda 2^{\lambda}} \sum_{m=1}^{2^{\lambda}} \sum_{\substack{n=1 \\ n \neq m}}^{2^{\lambda}} \Pr\left(\mathbf{x}_m \to \mathbf{x}_n\right) \chi\left(\mathbf{x}_m, \mathbf{x}_n\right), \tag{3.28}$$

where $\chi\left(\mathbf{x}_m, \mathbf{x}_n\right)$ is the number of different bits when comparing $\mathbf{x}_m$ and $\mathbf{x}_n$.

3.3.3 Simulation Results

In this section, the simulation results in terms of the bit error rate (BER) of the proposed IQSM and other recent SM schemes are presented and compared for various numbers of transmit antennas and spectral efficiency values versus the average SNR per receive antenna, which is defined as E_s/N_0. In all simulations, four receive antennas are employed. The relative error performance of different systems can be judged based on the minimum Euclidean distance normalized by the transmitted symbol energy, defined as:

$$d_{\min} = \min_{m,n} \sqrt{\frac{1}{E_s}} \|\mathbf{u}_m - \mathbf{u}_n\| \tag{3.29}$$

where $\mathbf{u}_m$ and $\mathbf{u}_n$ are two different realizations of transmit symbol vectors.

Figure 3.9 plots the theoretical upper bounds and simulation results of the proposed IQSM at 8, 10 and 12 bits/s/Hz. The figure clearly shows that the theoretical upper bounds are very tight to the actual performance at high SNR, which makes them useful in assessing the performance of the proposed IQSM for various antenna and modulation configurations.

Figure 3.10 shows the performance comparison between STBC-SM, QSM and IQSM at spectral efficiency values of 8 and 12 bits/s/Hz. All schemes use 4 transmit antennas. For the case of 8 bits/s/Hz, the proposed IQSM performs 4 dB better than QSM and 5 dB better than STBC-SM at BER $= 10^{-5}$. The better performance of IQSM over QSM can be examined by $d_{\min,\text{IQSM}} = 0.7071$ and $d_{\min,\text{QSM}} = 0.4472$. For STBC-SM, since this system

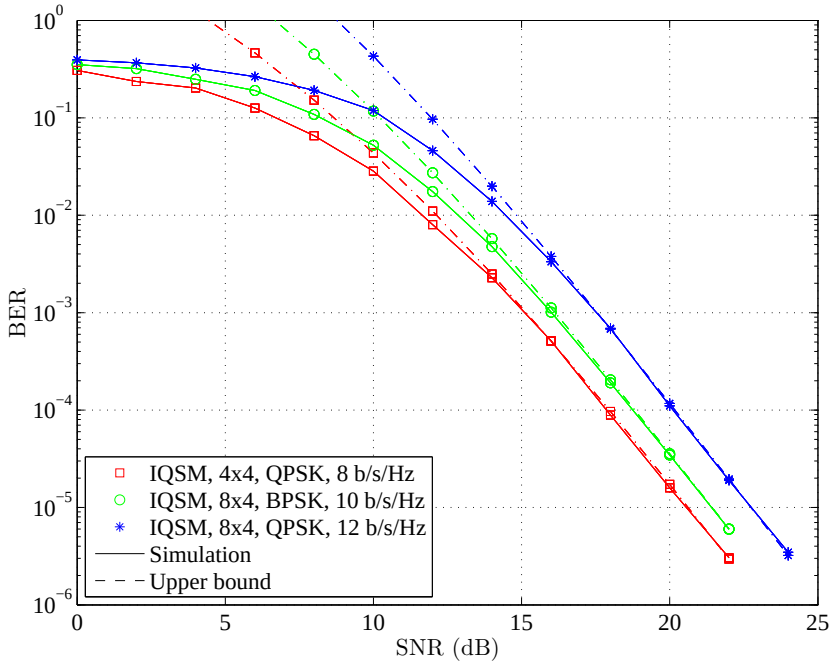

Figure 3.9 Simulation results and theoretical upper bounds of the bit error rate for the proposed IQSM.

achieves a higher transmit diversity gain than IQSM and QSM, the comparison in $d_{\min}$ is not meaningful. Recall that IQSM is the improved version of QSM by sending more than one constellation symbol and a larger number of indexing bits. Hence, at the same data rate, IQSM can use a lower-order constellation compared to the constellation used in QSM or STBC-SM. In this specific case, IQSM uses 2 symbols drawn from QPSK while QSM uses 16-QAM and STBC-SM uses 128-QAM. It is pointed out that, although the STBC-SM scheme achieves a transmit diversity order of two, the diversity gain only "kicks in" at high SNR, i.e., very low BER level, such as BER $< 10^{-8}$. At BER $= 10^{-5}$, the performance advantage of IQSM is even bigger at spectral efficiency of 12 bits/s/Hz. Specifically, IQSM yields a 8 dB better than QSM ($d_{\min,\mathrm{IQSM}} = 0.3162$, $d_{\min,\mathrm{QSM}} = 0.1085$) and IQSM is also superior than STBC-SM by 11 dB.

Further performance comparison of the proposed IQSM, STBC-SM and QSM is plotted in Figure 3.11 when the number of transmit antennas is increased to 8. As can be seen, at 10 bits/s/Hz, IQSM yields a 4 dB SNR gain compared to QSM ($d_{\min,\mathrm{IQSM}} = 0.7071$, $d_{\min,\mathrm{QSM}} = 0.4472$), and 7 dB SNR gain compared to STBC-SM. Compared to the 4-transmit

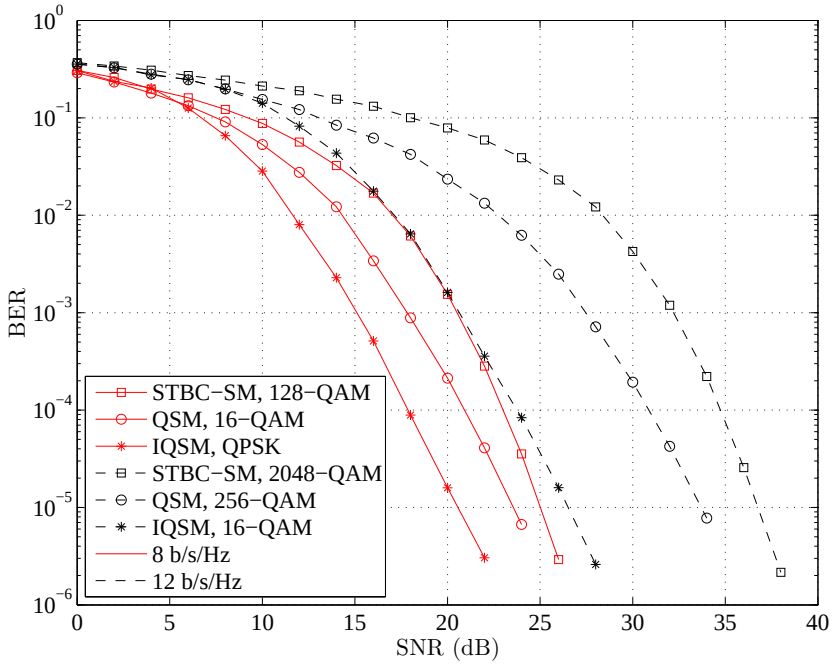

Figure 3.10 BER performance comparison of the proposed IQSM, STBC-SM and QSM at 8 bits/s/Hz and 12 bits/s/Hz using 4 transmit and 4 receive antennas (4×4 MIMO).

antenna configuration, with 8 transmit antenna, there are more bits that can be sent using the spatial dimension, i.e., by antenna indexing. This means that even lower-order constellations can be used in IQSM for the same spectral efficiency. For example, at 10 bits/s/Hz, while QSM still has to use 16-QAM constellation, IQSM can use BPSK instead of QPSK and the performance advantage of IQSM over QSM is about 6 dB SNR gain. It is pointed out that, since IQSM works with non-zero real and imaginary components of constellation points, the BPSK constellation is rotated with $\alpha = \pi/4$.

In Figure 3.12, the proposed IQSM is compared with spatial multiplexing (SMUX) and conventional SM. SMUX is a multiplexing transmission method to maximize the capacity gain. In SMUX, each transmit antenna is assigned to a fixed data stream (i.e., the number of transmit antennas is equal to the number of data streams). Each transmit antenna sends out one data stream independently from all other data streams and there is no antenna selection as in the SM-based schemes. At the spectral efficiency of 8 bits/s/Hz, SMUX is considered for two cases: (i) 2 transmit (2-Tx) antennas with 16-QAM, and (ii) 4 transmit antennas

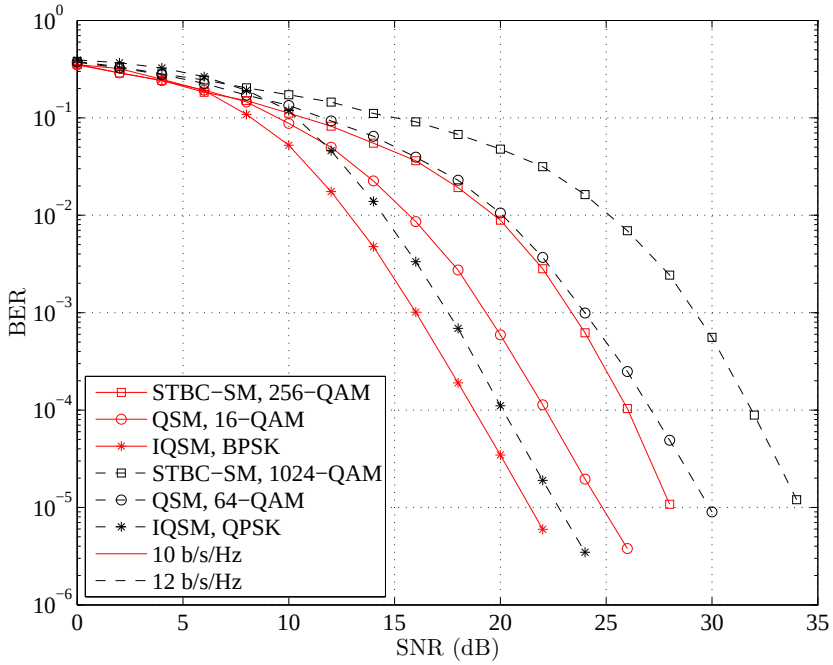

Figure 3.11 BER performance comparison of the proposed IQSM, STBC-SM and QSM at 10 bits/s/Hz and 12 bits/s/Hz using 8 transmit antennas and 4 receive antennas (8×4 MIMO).

(4-Tx) with QPSK. The proposed IQSM and conventional SM use 4 transmit antennas with QPSK and 64-QAM, respectively. It can be seen from Figure 3.12 that SMUX with 4 transmit antennas and QPSK has about 1 dB better than IQSM, while SMUX with 2 transmit antennas and 16-QAM is about 3 dB worse than IQSM. The minimum distance values of these schemes are found to be $d_{\text{min,IQSM}} = 0.7071$, $d_{\text{min,SMUX,4-Tx}} = 0.7071$ and $d_{\text{min,SMUX,2-Tx}} = 0.4472$. Even though SMUX with 4-Tx and IQSM have the same minimum distance, SMUX has a slight better performance because the number of transmit vector pairs having the distance d_{min} in SMUX is smaller than that in IQSM. It is pointed out that IQSM has about 5 dB SNR gain over the conventional SM, which has $d_{\text{min,SM}} = 0.3086$.

To deliver 12 bits/s/Hz, IQSM and conventional SM employ 8 transmit antennas and SMUX uses 3 or 4 transmit antennas. At BER $= 10^{-5}$, IQSM is still 2 dB better than SMUX with 4-Tx and 3 dB better than SMUX with 3-Tx. IQSM also yields a 10 dB SNR gain compared to the conventional SM. Again, these gains are verified with the minimum distances of $d_{\text{min,IQSM}} = 0.7071$, $d_{\text{min,SMUX,4-Tx}} = 0.4082$, $d_{\text{min,SMUX,3-Tx}} = 0.3651$ and $d_{\text{min,SM}} = 0.1101$.

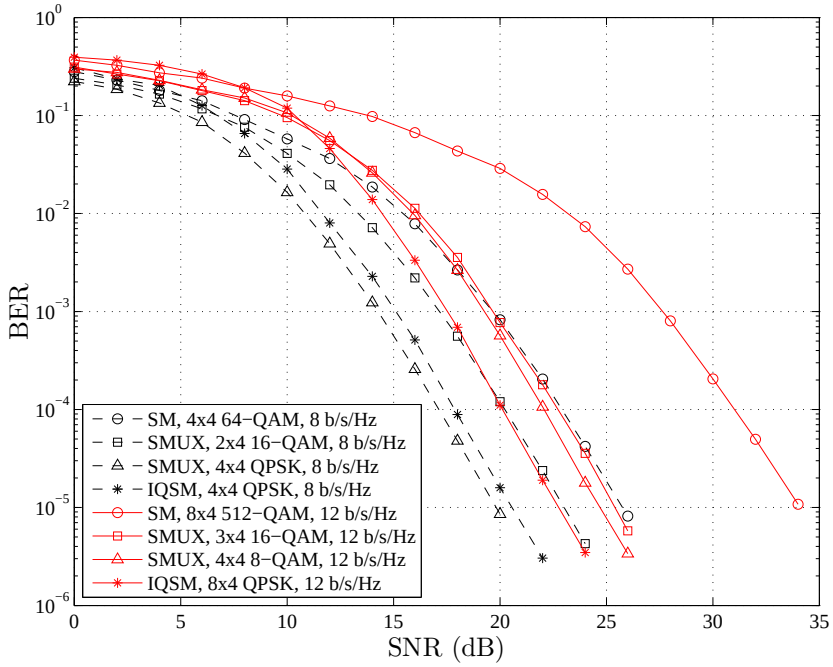

Figure 3.12 BER performance comparison of the proposed IQSM, SMUX and conventional SM at 8 bits/s/Hz and 12 bits/s/Hz.

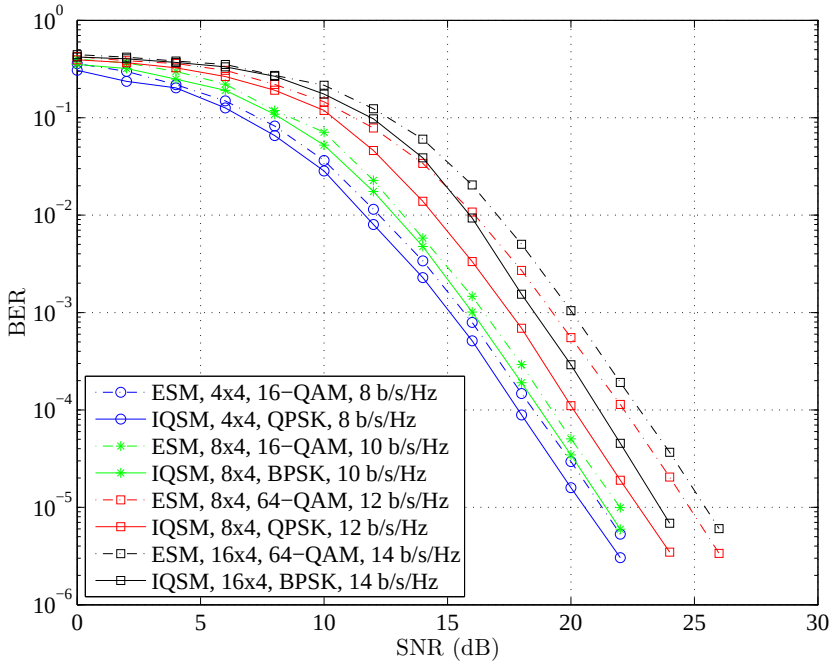

Figure 3.13 BER performance comparison of IQSM and enhanced SM (ESM) with different settings.

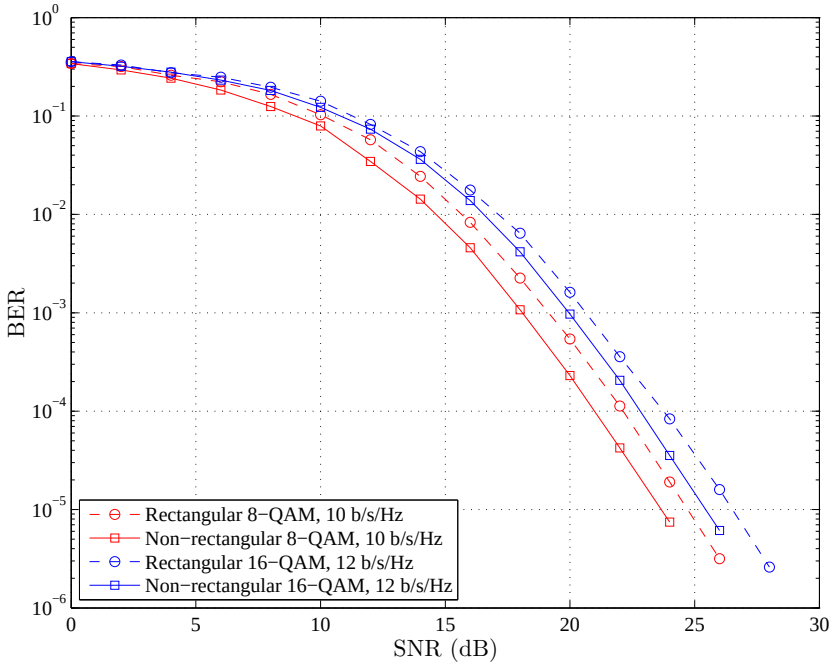

Figure 3.14 BER performance comparison of IQSM with rectangular and non-rectangular QAM constellations: 4×4 MIMO channel.

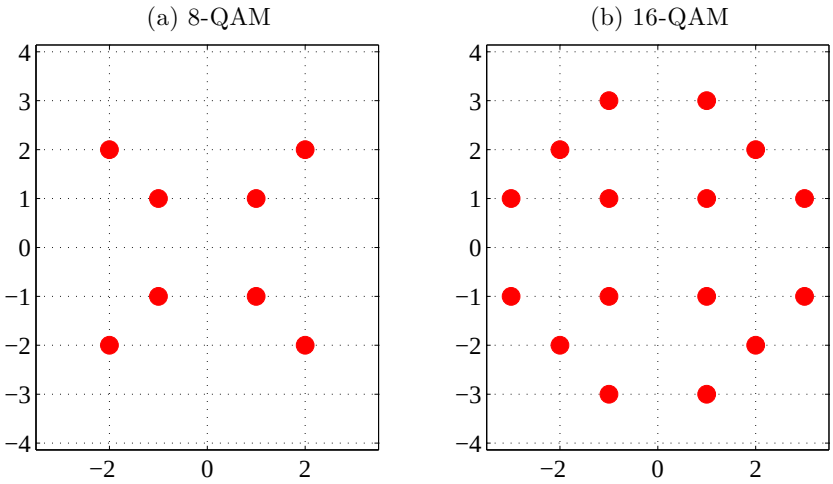

Figure 3.15 Non-rectangular 8-QAM and 16-QAM constellations.

From the simulation, it is noted that to increase the spectral efficiency, SMUX should be employed with a larger number of transmit antennas (i.e., number of data streams) rather than with a higher modulation order. However, when the number of transmit antennas is increased, the number of RF chains also increases. On the contrary, the maximum number of RF chains in the proposed IQSM is always 4.

As discussed before, reference [49] recently proposed the so called enhanced spatial modulation (ESM) employing multiple signal constellations. To date, this scheme is probably the most efficient in improving the spectral efficiency or error performance of SM. Figure 3.13 compares the BER performance of the proposed IQSM and ESM with different settings. It can be seen that, at 8 and 10 bits/s/Hz, the proposed IQSM is about 1 dB better than ESM. At a higher spectral efficiency, such as 12 and 14 bits/s/Hz, the proposed IQSM is clearly superior than ESM, by more than 2 dB. Again, the better performance of IQSM over ESM can be expected by examining the values of $d_{\min}$. For example, for the case of 14 bits/s/Hz using 16 transmit antennas, one can verify that $d_{\min,\text{IQSM}} = 0.7071$ and $d_{\min,\text{ESM}} = 0.3990$.

Finally, the performance of the proposed IQSM in Figure 3.14 for the case of 4×4 MIMO channel show that the non-rectangular 8-QAM and 16-QAM constellations in Figure 3.15 are more power efficient than the standard rectangular QAM. In particular, at BER $= 10^{-5}$, using the special 8-QAM and 16-QAM yields about 1 dB gain as compared to rectangular 8-QAM and 16-QAM, respectively. The minimum distances are found to be $d_{\min,\text{non-rect 8-QAM}} = 0.4472$, $d_{\min,\text{rect 8-QAM}} = 0.4082$, $d_{\min,\text{non-rect 16-QAM}} = 0.3651$ and $d_{\min,\text{rect 16-QAM}} = 0.3162$. This observation suggests that there is a significant benefit in optimizing QAM constellations for IQSM.

3.4 Constellation Design for Quadrature Spatial Modulation

Recently, the optimal constellation design for a SM system is investigated in [3] by minimizing the error probability. It is demonstrated in [3] that the error performance of a SM system depends on two terms: (i) the Euclidean distances among APM symbols, and (ii) the energies of APM symbols [50, 51]. While the first term is related to the error events associated with the conventional APM constellation, the second term is related to the error events associated with the spatial domain, which is the new feature in SM. In addition, the authors in [3] show that a phase shift keying (PSK) constellation is the optimal constellation for SM at very large number of transmit antennas, whereas increasing the number of receive antennas increases the dominance of the Euclidean distance term over the energy term. In [52], a star-QAM constellation is proposed as a good candidate for SM and a twin-ring

constellation design is introduced to minimize the error probability.

Quadrature spatial modulation (QSM) was recently proposed [48] to enhance the spectral efficiency of SM. In QSM, the in-phase and quadrature components are separately transmitted on two different layers of transmit antennas. Thus, the number of bits sent by antenna index in QSM doubles that in SM. With different strategies for selecting active antennas between QSM and SM, it is expected that the optimal constellation design for QSM is not the same as that for SM. To the best of our knowledge, to date there is no study on optimizing QSM constellations. To fill this gap, this section first analyzes an upper bound on the error probability of QSM under the maximum likelihood (ML) detection to understand the main factors that affect the error performance in QSM. The analysis shows that the error probability of QSM not only includes the Euclidean distance term and the energy term (the same as in the case of SM), but also two new terms that are related to the in-phase and quadrature components of APM symbols. These extra two terms arise as a consequence of the unique property of QSM compared to SM. Based on the obtained analysis, an optimization problem is formulated to find the optimal constellations for QSM under different numbers of transmit antennas. It is shown that with a very large number of transmit antennas the optimal constellations of QSM converge to a QPSK constellation. To verify the advantage of the optimal constellations over other constellations, extensive performance comparison is carried out by computer simulation.

3.4.1 System Model and Performance Analysis

Consider a MIMO system that is equipped with N_t transmit and N_r receive antennas. The equivalent baseband input/output signal model is given by

$$\mathbf{y} = \sqrt{E_s}\mathbf{H}\mathbf{x} + \mathbf{n} \tag{3.30}$$

In the above expression, $\mathbf{x}$ is an $N_t \times 1$ transmit symbol vector comprising of nonzero elements (corresponding to active antennas) and zero elements (corresponding to idle antennas), $\mathbf{y}$ is an $N_r \times 1$ received signal vector, while $\mathbf{n}$ is an $N_r \times 1$ noise vector whose entries are independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and variance N_0. The $N_r \times N_t$ matrix $\mathbf{H}$ is the channel matrix, whose entries are i.i.d

complex Gaussian random variables with zero mean and unit variance. This means that the channel is flat Rayleigh fading. The element on the mth row and nth column of $\mathbf{H}$ is denoted as $h_{m,n}$, which represents the complex channel path gain between the nth transmit antenna and the mth receive antenna. Each transmitted vector $\mathbf{x}$ has a unit norm and the constant E_s is interpreted as the average transmitted energy per each transmit symbol vector $\mathbf{x}$.

In QSM, the real and imaginary components of a constellation symbol are transmitted by two different layers of antennas (in-phase and quadrature dimensions). Therefore, the received signal vector can be re-written in the following form:

$$\mathbf{y} = \sqrt{E_s}(\mathbf{h}_{n_\Re} x_{\ell_\Re} + j\mathbf{h}_{n_\Im} x_{\ell_\Im}) + \mathbf{n}, \quad 1 \leq n_\Re, n_\Im \leq N_t, \tag{3.31}$$

where $x_\ell = x_{\ell_\Re} + j x_{\ell_\Im}$ is a point drawn from a constellation with size M ($1 \leq \ell \leq M$) and $\mathbf{h}_{n_\Re}$ is the $n_\Re$th column of $\mathbf{H}$. The optimal ML detection for QSM is expressed as:

$$[\tilde{n}_\Re, \tilde{n}_\Im, \tilde{\ell}] = \arg \min_{n_\Re, n_\Im, \ell} \left\| \mathbf{y} - \sqrt{E_s}(\mathbf{h}_{n_\Re} x_{\ell_\Re} + j\mathbf{h}_{n_\Im} x_{\ell_\Im}) \right\|^2 \tag{3.32}$$

With QSM, a total of $K = N_t^2 M$ different signal vectors $\mathbf{x}$ can be generated, i.e., the signal space is composed of $N_t^2 M$ codewords. Given the channel matrix $\mathbf{H}$, an upper bound on the codeword error probability of QSM under the ML detection rule can be expressed as

$$\begin{aligned} \Pr(\text{error} \mid \mathbf{H}) \leq \frac{1}{K} \sum_{n_\Re=1}^{N_t} \sum_{n_\Im=1}^{N_t} \sum_{\ell=1}^{M} \sum_{n'_\Re=1}^{N_t} \sum_{n'_\Im=1}^{N_t} \sum_{\ell'=1}^{M} \\ Q\left(\sqrt{\frac{E_s}{2N_0}} \left\| \mathbf{h}_{n_\Re} x_{\ell_\Re} + j\mathbf{h}_{n_\Im} x_{\ell_\Im} - \mathbf{h}_{n'_\Re} x_{\ell'_\Re} - j\mathbf{h}_{n'_\Im} x_{\ell'_\Im} \right\| \right), \end{aligned} \tag{3.33}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right)$ is the Q-function. The above union bound can be re-arranged as in (3.34) at the bottom of the next page. Taking the expectation from both sides of (3.34) and using the same approximation method as in [6, 50], the union bound for the average codeword error probability can be generally expressed as

$$\begin{aligned} \Pr(\text{error}) \approx \begin{pmatrix} 2N_r - 1 \\ N_r \end{pmatrix} \frac{(E_s/N_0)^{-N_r}}{M} \times \\ \Big\{ (N_t - 1)^2 \Omega_1 + (N_t - 1)\Omega_2 + (N_t - 1)\Omega_3 + \Omega_4 \Big\}, \end{aligned} \tag{3.35}$$

$$\Omega_1 = \sum_{\ell=1}^{M}\sum_{\ell'=1}^{M}\left[x_{\ell_{\Re}}^2 + x_{\ell_{\Im}}^2 + x_{\ell'_{\Re}}^2 + x_{\ell'_{\Im}}^2\right]^{-N_r} \tag{3.36}$$

$$\Omega_2 = \sum_{\ell=1}^{M}\sum_{\ell'=1}^{M}\left[(x_{\ell_{\Re}} - x_{\ell'_{\Re}})^2 + x_{\ell_{\Im}}^2 + x_{\ell'_{\Im}}^2\right]^{-N_r} \tag{3.37}$$

$$\Omega_3 = \sum_{\ell=1}^{M}\sum_{\ell'=1}^{M}\left[x_{\ell_{\Re}}^2 + x_{\ell'_{\Re}}^2 + (x_{\ell_{\Im}} - x_{\ell'_{\Im}})^2\right]^{-N_r} \tag{3.38}$$

$$\Omega_4 = \sum_{\ell=1}^{M}\sum_{\ell'=1,\ell'\neq\ell}^{M}\left[(x_{\ell_{\Re}} - x_{\ell'_{\Re}})^2 + (x_{\ell_{\Im}} - x_{\ell'_{\Im}})^2\right]^{-N_r} \tag{3.39}$$

It is pointed out that the two terms Ω_1 and Ω_4 in (3.35) are also present in the upper bound expression for the case of an SM system, in which Ω_1 is simply the sum of the energies of APM symbols and Ω_4 is the sum of the squared Euclidean distances among APM symbols. The two additional terms, Ω_2 and Ω_3, in (3.35) are unique for QSM and related to the in-phase and quadrature components of APM symbols.

$$\begin{aligned}
\Pr(\text{error} \mid \mathbf{H}) \;\leq\; & \frac{1}{K}\sum_{\substack{n_{\Re}=n'_{\Re}\\ n_{\Im}=n'_{\Im}}}^{N_t}\sum_{\ell=1}^{M}\sum_{\substack{\ell'=1\\ \ell\neq\ell'}}^{M} Q\left(\sqrt{\frac{E_s}{2N_0}}\left\|(x_{\ell_{\Re}} - x_{\ell'_{\Re}})\mathbf{h}_{n_{\Re}} + j(x_{\ell_{\Im}} - x_{\ell'_{\Im}})\mathbf{h}_{n_{\Im}}\right\|\right) \\
& + \frac{1}{K}\sum_{\substack{n_{\Re}\neq n'_{\Re}\\ n_{\Im}=n'_{\Im}}}^{N_t}\sum_{\ell=1}^{M}\sum_{\ell'=1}^{M} Q\left(\sqrt{\frac{E_s}{2N_0}}\left\|\mathbf{h}_{n_{\Re}}x_{\ell_{\Re}} - \mathbf{h}_{n'_{\Re}}x_{\ell'_{\Re}} + j(x_{\ell_{\Im}} - x_{\ell'_{\Im}})\mathbf{h}_{n_{\Im}}\right\|\right) \\
& + \frac{1}{K}\sum_{\substack{n_{\Re}=n'_{\Re}\\ n_{\Im}\neq n'_{\Im}}}^{N_t}\sum_{\ell=1}^{M}\sum_{\ell'=1}^{M} Q\left(\sqrt{\frac{E_s}{2N_0}}\left\|(x_{\ell_{\Re}} - x_{\ell'_{\Re}})\mathbf{h}_{n_{\Re}} + j\mathbf{h}_{n_{\Im}}x_{\ell_{\Im}} - j\mathbf{h}_{n'_{\Im}}x_{\ell'_{\Im}})\right\|\right) \\
& + \frac{1}{K}\sum_{\substack{n_{\Re}\neq n'_{\Re}\\ n_{\Im}\neq n'_{\Im}}}^{N_t}\sum_{\ell=1}^{M}\sum_{\ell'=1}^{M} Q\left(\sqrt{\frac{E_s}{2N_0}}\left\|\mathbf{h}_{n_{\Re}}x_{\ell_{\Re}} + j\mathbf{h}_{n_{\Im}}x_{\ell_{\Im}} - \mathbf{h}_{n'_{\Re}}x_{\ell'_{\Re}} - j\mathbf{h}_{n'_{\Im}}x_{\ell'_{\Im}}\right\|\right)
\end{aligned} \tag{3.34}$$

3.4.2 QSM Constellation Design

Since the union (upper) bound becomes tight in the high signal-to-noise ratio (SNR) region, it is common to design a transmission scheme to minimize the upper bound of the average error probability. Given the expression of the bound in (3.35), the problem of optimizing a QSM constellation for fixed N_t, N_r and M can be formulated as follows:

$$\begin{aligned} \underset{\substack{x_{\ell_\Re}, x_{\ell_\Im}(\ell=1,\ldots,M) \\ -\sqrt{M} \le x_{\ell_\Re}, x_{\ell_\Im} \le \sqrt{M}}}{\arg\min} \quad & (N_t-1)^2\Omega_1 + (N_t-1)(\Omega_2+\Omega_3) + \Omega_4 \\ \text{subject to } & \sum_{\ell=1}^{M} \left(x_{\ell_\Re}^2 + x_{\ell_\Im}^2\right) = M. \end{aligned} \tag{3.40}$$

Unfortunately, finding a closed-form solution for the above optimization appears very difficult. Instead, Matlab Optimization Toolbox is used to obtain the solution numerically. The obtained constellations for different combinations of N_t, N_r and M are plotted in Figure 3.16. Observe that when the number of transmit antennas, N_t, is large, M APM symbols are evenly and symmetrically distributed over four clusters. In fact, when N_t approaches ∞, the following analysis shows that the four clusters converge to a QPSK-like constellation.

Observe that when N_t is very large, the term $(N_t-1)^2\Omega_1$ is most dominant, followed by $(N_t-1)(\Omega_2+\Omega_3)$, whereas Ω_4 can be ignored. Thus it is reasonable to minimize Ω_1 first, then $\Omega_2+\Omega_3$, and lastly Ω_4.

First, under the constrain in (3.40), it has been shown in [3] that in order to minimize Ω_1, all the APM symbols have equal (unit) energy, i.e., $x_{\ell_\Re}^2 + x_{\ell_\Im}^2 = 1$ for $\forall \ell$. Next, the optimization problem to minimize $\Omega_2+\Omega_3$ can be formulated as follows:

$$\begin{aligned} \underset{\substack{x_{\ell_\Re}, x_{\ell_\Im}(\ell=1,\ldots,M) \\ -1 \le x_{\ell_\Re}, x_{\ell_\Im} \le 1}}{\arg\min} \quad & f = \sum_{\ell=1}^{M}\sum_{\ell'=1}^{M} f_{(\ell,\ell')} \\ \text{subject to } \quad & x_{\ell_\Re}^2 + x_{\ell_\Im}^2 = 1, \quad 1 \le \ell \le M \end{aligned} \tag{3.41}$$

where $f_{(\ell,\ell')} = (1 - x_{\ell_\Re} x_{\ell'_\Re})^{-N_r} + (1 - x_{\ell_\Im} x_{\ell'_\Im})^{-N_r}$. To get the solution for (3.41), first it can be shown that $f_{(\ell,\ell)} = (1 - x_{\ell_\Re}^2)^{-N_r} + (1 - x_{\ell_\Im}^2)^{-N_r}$ is minimized when $x_{\ell_\Re}^2 = x_{\ell_\Im}^2 = 1/2$. Furthermore, $f_{(\ell,\ell)}$ will be infinity for any of the following four "singular" locations on the

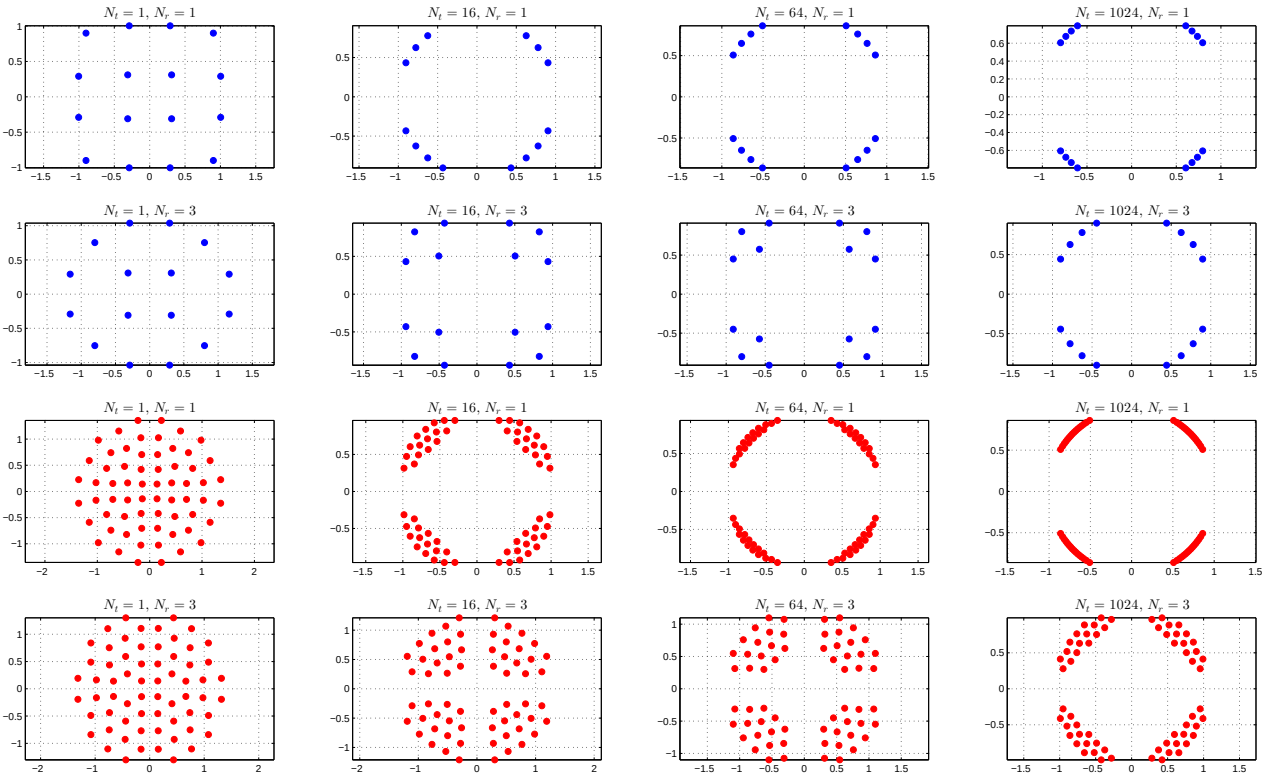

Figure 3.16 Optimal QSM constellations for $M = 16$ (blue) and $M = 64$ (red) and different combinations of N_t and N_r.

unit circle: $(0, 1)$, $(1, 0)$, $(0, -1)$, $(-1, 0)$. Next, observe that for any signal pair (ℓ, ℓ'), the corresponding function $f_{(\ell,\ell')}$ would be minimized if the two signals are in opposite quadrants of the unit circle (so that both products $x_{\ell_\Re} x_{\ell'_\Re}$ and $x_{\ell_\Im} x_{\ell'_\Im}$ are negative). This means that for the function f in (3.41) to be minimized, there must be a maximum number of signal pairs that are in opposite quadrants. It turns out that the solution of (3.41) distributes all M signals equally either over two opposite quadrants or over all four quadrants at locations $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$, i.e., the locations of standard QPSK signal points.

It is pointed out that the uncertainty (ambiguity) of the solution of (3.41) is a direct consequence of assuming N_t to be very large and ignoring the term Ω_4 in (3.40). However, for the practical case of having a large, but limited number of transmit antennas, the minimization of Ω_4 should be considered. Since Ω_4 is inversely proportional to the Euclidian

distances among the signal points, one expects that all the signal points are distinct, i.e., no two signal points overlap. This means that all M signal points are clustered around the four nominal QPSK points instead of lying exactly on these points as when the term Ω_4 is ignored. The distribution of M signal points among the four clusters that minimizes Ω_4 is addressed next.

Express $\Omega_4 = \sum_{\ell=1}^{M} \sum_{\ell'=1}^{M} \left(d_{(\ell,\ell')}^2 \right)^{-N_r}$ where $d_{(\ell,\ell')}$ is the Euclidean distance between x_ℓ and $x_{\ell'}$ ($\ell \neq \ell'$). Consider the general case that the numbers of constellation points that are clustered around the four nominal QPSK points are M_1, M_2, M_3 and M_4, as illustrated in Figure 3.17. The optimization problem is

$$\begin{aligned} \underset{\substack{M_1,M_2,M_3,M_4 \\ 0 \leq M_i \leq M}}{\arg\min} \quad & \Omega_4 = \sum_{\ell=1}^{M} \sum_{\ell'=1,\ell' \neq \ell}^{M} \left(d_{(\ell,\ell')}^2 \right)^{-N_r} \\ \text{subject to} \quad & M_1 + M_2 + M_3 + M_4 = M. \end{aligned} \tag{3.42}$$

To find the solution for the above problem, let $\epsilon > 0$ be the maximum distance of any signal point to the nominal QPSK point within any of the four clusters and d_0 be the minimum Euclidean distance among the four nominal QPSK points (see Figure 3.17). Under the assumption that $\epsilon \ll d_0$ for very large N_t, $d_{(\ell,\ell')}^{-2N_r}$ takes on one of three values: $\alpha_0 = \epsilon^{-2N_r}$, $\alpha_1 = d_0^{-2N_r}$ and $\alpha_2 = (\sqrt{2}d_0)^{-2N_r}$, where $\alpha_0 \gg \alpha_1 > \alpha_2 > 0$. Then, the function Ω_4 can be written in terms of variables M_1, M_2, M_3, M_4 as follows:

$$\begin{aligned} \Omega_4 \approx & [M_2\alpha_1 + M_4\alpha_1 + M_3\alpha_2 + (M_1 - 1)\alpha_0]M_1 \\ & + [M_1\alpha_1 + M_3\alpha_1 + M_4\alpha_2 + (M_2 - 1)\alpha_0]M_2 \\ & + [M_2\alpha_1 + M_4\alpha_1 + M_1\alpha_2 + (M_3 - 1)\alpha_0]M_3 \\ & + [M_1\alpha_1 + M_3\alpha_1 + M_2\alpha_2 + (M_4 - 1)\alpha_0]M_4 \end{aligned} \tag{3.43}$$

Using the constraint $M_1 + M_2 + M_3 + M_4 = M$, the objective function in (3.43) can be simplified to $\hat{f} = (M_1^2 + M_2^2 + M_3^2 + M_4^2)\beta_1 - 2(M_1M_3 + M_2M_4)\beta_2$, where $\beta_1 = (\alpha_0 - \alpha_1)$, $\beta_2 = (\alpha_1 - \alpha_2)$ and $\beta_1 \gg \beta_2 > 0$. Taking the derivative of the Lagrangian function with respect to M_1, M_2, M_3, M_4 and setting the results to zero gives $M_1 = M_3$ and $M_2 = M_4$. Then substitute (M_1, M_2) for (M_3, M_4) in $\hat{f}$ to have a new optimization problem: minimizing

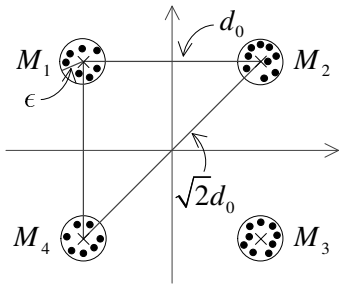

Figure 3.17 Distribution of M signal points over four clusters centered at nominal QPSK points and relevant distances.

$(M_1^2 + M_2^2)(\beta_1 - \beta_2)$ under the constraint $M_1 + M_2 = M/2$. The solution for this simple optimization problem is $M_1 = M_2 = M/4$.

To summarize, for the case of having a large but limited number of transmit antennas, the structure of the optimal QSM constellation is that all M signal points are equally divided over four clusters that are centered around the four nominal QPSK points. Such solution structure can be seen in Figure 3.16 for $N_t = 16, 64, 1024$.

3.4.3 Simulation Results and Comparisons

In this section, simulation results in terms of the bit error rate (BER) of QSM employing the proposed constellations and other constellations such as rectangular QAM, PSK and constellations optimally designed for SM in [3] are presented. The results are compared for various numbers of transmit antennas and spectral efficiency values versus the average SNR per receive antenna, which is defined as E_s/N_0. In all simulations, $N_r = 3$ receive antennas are employed.

In Figure 3.18, QSM using the proposed constellation is compared with QSM using rectangular QAM and PSK at spectral efficiency values of 12 b/s/Hz and 16 b/s/Hz. All constellations have the same size of $M = 16$. To achieve 12 b/s/Hz, all three QSM systems employ $N_t = 16$, while for 16 b/s/Hz, $N_t = 64$. It is pointed out that, since a QSM system works better for constellations whose in-phase and quadrature components are non zero, the PSK constellations need to be rotated. At 12 b/s/Hz and 16 b/s/Hz, it can be seen that the proposed constellations yield 2 dB better than the rectangular 16-QAM, and 15 dB better

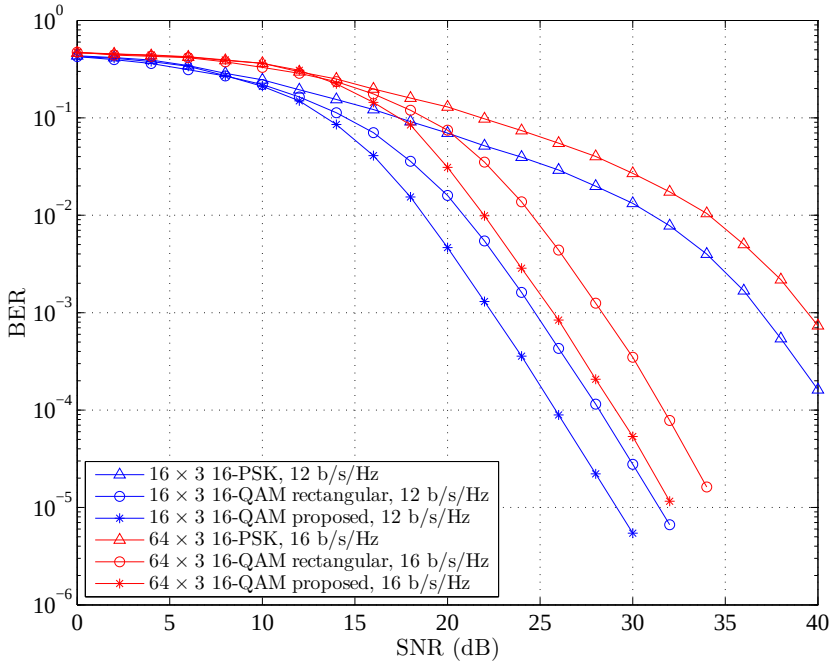

Figure 3.18 Performance of QSM with phase shift keying, rectangular QAM and the proposed constellations at 12 b/s/Hz and 16 b/s/Hz.

than the rotated PSK.

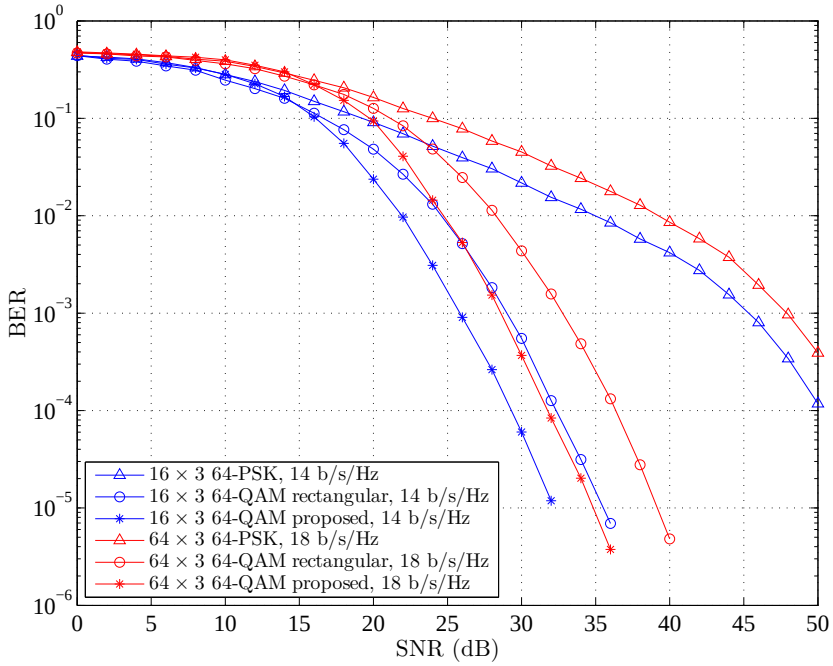

Figure 3.19 Performance of QSM with phase shift keying, rectangular QAM and the proposed constellations at 14 b/s/Hz and 18 b/s/Hz.

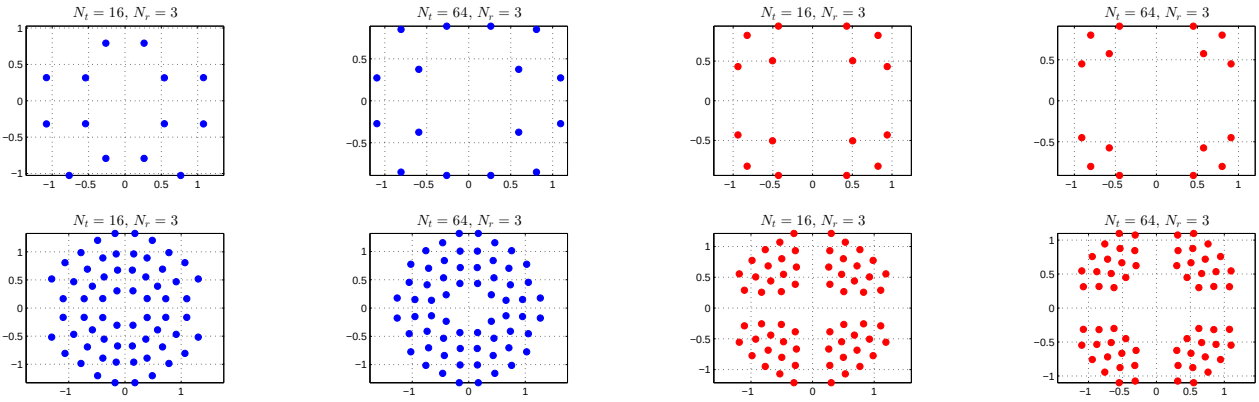

Figure 3.20 Comparison of proposed constellations (designed for QSM) and constellations designed for SM in [3].

Similar performance comparison is presented in Figure 3.19, but for 14 b/s/Hz and 18 b/s/Hz. To achieve these spectral efficiency values, the constellation size is increased to $M = 64$, while $N_t = 16$ and $N_t = 64$ transmit antennas are employed, respectively. It can be seen that at higher transmission rates, the proposed 64-ary constellations are still much better than the rectangular 64-QAM and 64-PSK. In particular, at 14 b/s/Hz the proposed constellation is 2 dB better than the rectangular 64-QAM and 20 dB better than the rotated 64-PSK. At 18 b/s/Hz, the SNR gain of the proposed 64-ary constellation increases to 4 dB compared to the rectangular 64-QAM.

Finally, Figure 3.20 compares the geometry of constellations designed for QSM in this thesis (shown in red) and the constellation designed for SM in [3] (shown in blue) for different combinations of the numbers of transmit and receive antennas. The top panel shows 16-ary constellations, while the bottom panel shows 64-ary constellations. For the same setting (N_t, N_r and M), the clear difference of constellations designed for QSM and SM is again due to the fact that a signal constellation impacts the error performance of QSM and SM systems differently. It can be seen from Figure 3.21 that, when applied to QSM systems, the proposed constellations (as designed in this thesis) consistently outperform the constellations designed for SM in [3] by 1 to 2 dB.

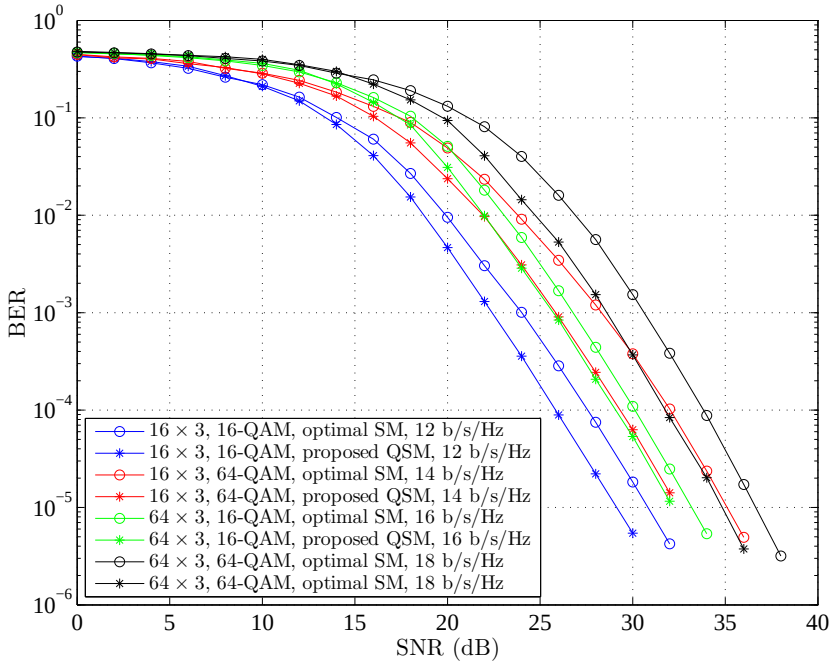

Figure 3.21 Performance comparison between the proposed constellations and constellations designed for SM in [3].

3.5 Summary

This chapter has developed new transmission techniques for SM to improve the spectral efficiency over frequency-flat fading channels. A novel transmission scheme was developed by combining SM and a high-rate space-time block code to achieve the transmit diversity order of two at high spectral efficiency. For a SSK system, a new modulation scheme was proposed by multiplexing in-phase and quadrature GSSK streams to double the transmission rate of GSSK. To generalize the second scheme for a SM system (i.e., constellation symbols are also used), IQSM was proposed in which the number of spatial constellation points for each antenna dimension is increased and an additional modulation symbol is sent in the same symbol duration. Finally, an optimal constellation design for QSM was presented to minimize the average error probability.

4. Spatial Modulation for Frequency-Selective Fading Channels

This chapter investigates the applications of SM in frequency-selective fading channels. First, a novel scheme, called spatial modulation OFDM with grouped linear constellation precoding (SM-OFDM-GLCP), is presented where subcarriers are divided in groups and information symbols are spread across subcarriers in each group to maximize the diversity and coding gains. The technique of SM is then introduced to select an antenna to transmit OFDM symbols of the group. Performance analysis and numerical results are presented to demonstrate the diversity and coding gains of the proposed scheme. Next, a comparison between IM-based OFDM and precoded-OFDM with multiple constellations is carried out to provide insights on the advantages and disadvantages of IM-based OFDM systems.

4.1 Spatial Modulation for OFDM with Linear Constellation Precoding

As explained in Chapter 2, orthogonal frequency-division multiplexing (OFDM) is a multi-carrier transmission method which is commonly used for frequency-selective fading channels due to its ability to transform a wideband system into a multiple narrowband systems. In classical OFDM, the detection process can be implemented on each subcarrier by a simple zero-forcing equalizer. However, classical OFDM does not achieve the inherent multipath diversity gain because each information symbol is not spread over independently-faded subcarriers. A linear constellation precoding (LCP) method is proposed in [53, 54] to improve the performance of a classical OFDM system. This method, called grouped-LCP OFDM (OFDM-GLCP), implements both subcarrier grouping and symbol spreading

over independently-faded subcarriers. With OFDM-GLCP the diversity and coding gains can be achieved with acceptable detection complexity. Another approach to improve the diversity gain of classical OFDM is OFDM with index modulation (OFDM-IM) [35]. As with OFDM-GLCP, in OFDM-IM, the total number of subcarriers is divided into groups and index modulation is applied in each group to select active subcarriers for transmitting information symbols.

This part of research investigates a scheme called spatial modulation OFDM-GLCP (SM-OFDM-GLCP) where subcarriers are divided in groups and information symbols are spread across subcarriers in each group to maximize the diversity and coding gains[1]. Furthermore, for each subcarrier group, spatial modulation is employed to select an antenna to transmit OFDM symbols of the group. By using spatial modulation for each subcarrier group, the advantage of spatial modulation in traditional MIMO can be exploited in MIMO-OFDM, i.e., avoiding ICI between symbols in the same subcarriers at different antennas. Using spatial modulation also increases the transmission rate of SM-OFDM-GLCP as compared to OFDM-GLCP since more information bits are conveyed by antenna indices. Performance analysis is carried out to determine the diversity and coding gains of the proposed scheme. Numerical results and comparison with existing schemes such as SM-OFDM and OFDM-GLCP demonstrate the efficiency of the proposed SM-OFDM-GLCP.

4.1.1 System Model

The transmitter of the SM-OFDM-GLCP system investigated in this research is illustrated in Figure 4.1 with N_t transmit antennas. There are N subcarriers in total, which are divided into G equal-size groups, each group having $F = N/G$ subcarriers. The numbers of groups and subcarriers in each group are identical. Furthermore, subcarrier grouping is done in an interleaved fashion so that the the group size is as small as possible (to reduce detection complexity) while subcarriers in each group are least correlated [53, 56]. A condition to achieve the full diversity gain is that the group size F is not less than the number of effective resolvable paths L of the multipath channel, i.e., $F \geq L$.

[1]The contributions in this section are presented in [55].

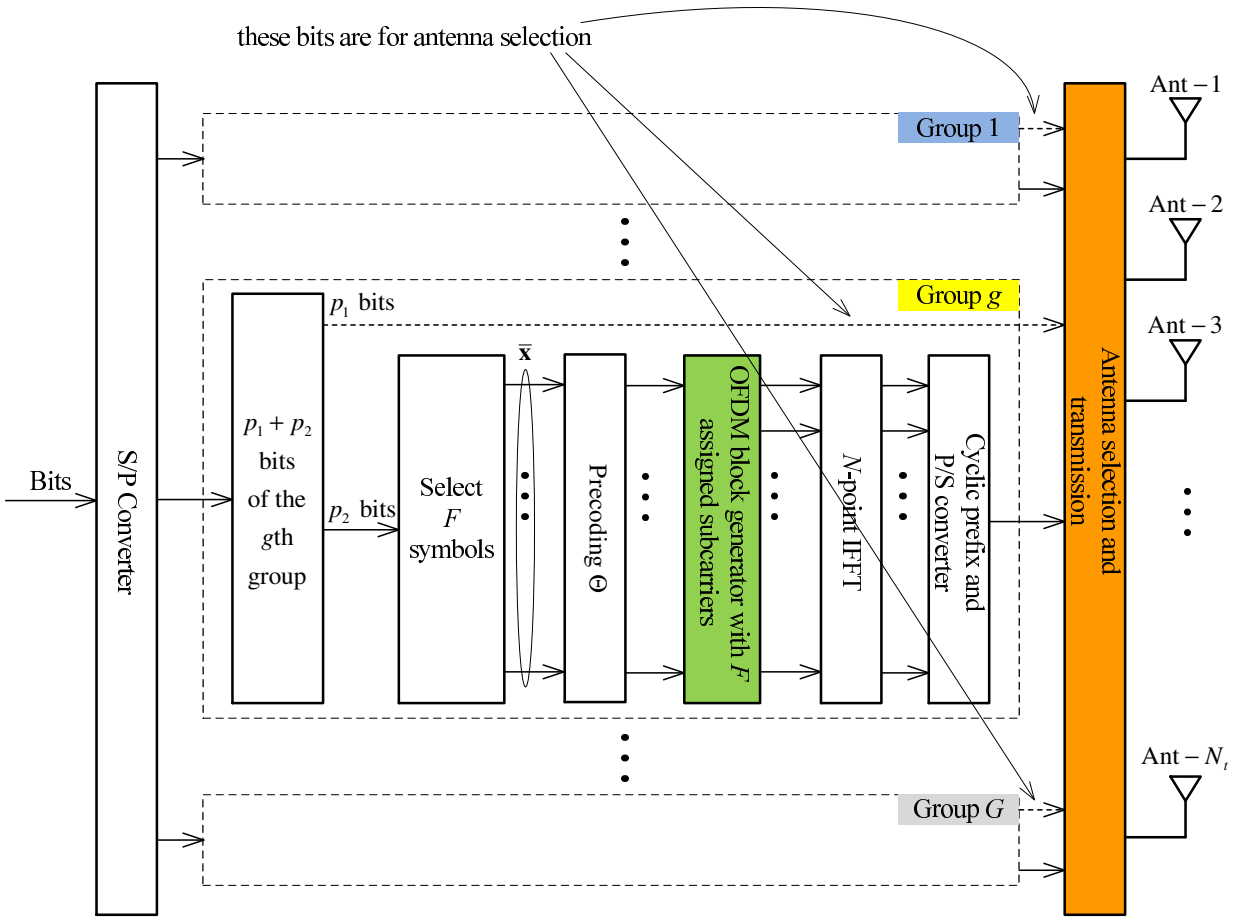

Figure 4.1 Transmitter of the proposed SM-OFDM-GLCP system.

The input bit sequence is organized into G groups as follows:

Group 1	Group 2	$\cdots$	Group G
p_1 bits	p_1 bits	$\cdots$	p_1 bits
p_2 bits	p_2 bits	$\cdots$	p_2 bits

Each group has $(p_1 + p_2)$ bits where p_1 bits are used for selecting a transmit antenna while p_2 bits determine F information symbols which are carried by F subcarriers of the group. These F information symbols form a column vector which is then multiplied with a rotation matrix, i.e., a precoder, Θ in order to achieve the maximum diversity and coding gains. The precoded symbols are mapped to appropriate subcarriers that are assigned to each group. It is pointed out that, while the subcarrier groups are disjoint and that all the N subcarriers are activated by the system, which transmit antenna is *effectively* used for a

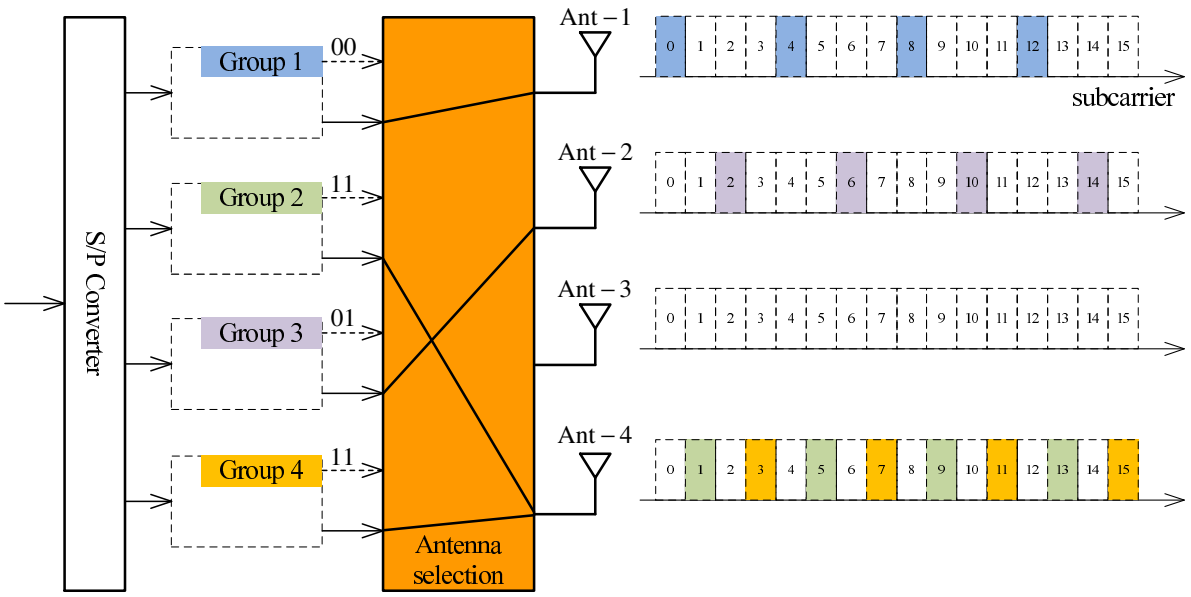

Figure 4.2 Illustration of the proposed SM-OFDM-GLCP scheme with $N_t = 4$ transmit antennas, $G = 4$ groups and $N = 16$ subcarriers.

given subcarrier group is determined by p_1 bits associated to each subcarrier group.

To illustrate the proposed scheme clearer, an example for the case of $N_t = 4$ transmit antennas, $G = 4$ groups and $N = 16$ subcarriers is provided in Figure 4.2. In this example the $p_1 = 2$ bits associated with the four groups are 00, 11, 01 and 11. As can be seen from the figure, with interleaved subcarrier grouping, the precoded symbols in group 1 are effectively sent over carrier numbers 0, 4, 8, 12 and over the first antenna, while the precoded symbols in group 2 are effectively sent over carrier numbers 1, 5, 9, 13 and over the fourth antenna. Likewise, the precoded symbols in group 3 are effectively sent over carrier numbers 2, 6, 10, 14 and over the second antenna, while the precoded symbols in group 4 are effectively sent over carrier numbers 3, 7, 11, 15 and over the fourth antenna. Note that in this specific example, antenna 3 is never effectively activated, while antenna 4 is effectively activated to send precoded symbols in both groups 2 and 4. It is pointed out that, from the time-domain perspective, all antennas in the proposed system are simultaneously activated.

It is clear that, under the maximum likelihood (ML) receiver, the error performance is identical for information bits sent over different groups. As such, it is sufficient to investigate the performance of any group to determine the overall system performance. In the following

analysis, the index of a subgroup is dropped for simplicity. Furthermore, in order to focus on the transmit diversity gain achieved by SM-OFDM-GLCP and to simplify the analysis, the number of receive antennas is limited to $N_r = 1$ in this thesis.

In the receiver, after the CP is removed from the time-domain samples, a block of N time samples are processed by an N-point DFT to obtain N frequency-domain samples. Focusing on any subcarrier group of interest, the input-output relationship in the frequency domain over F subcarriers of such a group is expressed as:

$$\mathbf{y} = [y(1), y(2), \cdots, y(F)]^T = \mathbf{X}\tilde{\mathbf{h}} + \mathbf{w} \tag{4.1}$$

where $\mathbf{w} = [w(1), w(2), \cdots, w(F)]^T$ represents additive white Gaussian noise (AWGN), whose elements are independent and identically distributed (i.i.d) zero-mean complex Gaussian random variables with variance N_0. In the above expression, the $F \times (F \cdot N_t)$ transmit data matrix $\mathbf{X}$ has the following form:

$$\mathbf{X} = [\mathbf{0}_{F\times F}, \mathbf{0}_{F\times F}, \cdots, \mathbf{X}_q, \cdots, \mathbf{0}_{F\times F}]\,. \tag{4.2}$$

Thus, $\mathbf{X}$ is built from N_t $F \times F$ sub-matrices in which there is only one non-zero sub-matrix $\mathbf{X}_q = \mathrm{diag}(\Theta\bar{\mathbf{x}})$. This non-zero sub-matrix is a $F \times F$ diagonal matrix, where $\bar{\mathbf{x}}$ is the $F \times 1$ symbol vector determined by p_2 bits and Θ is a $F \times F$ rotation matrix. The rotation matrix is designed with the unified algebraic construction method, LCP-A, as in [53]. The position of the sub-matrix $\mathbf{X}_q$ in $\mathbf{X}$ is determined by p_1 bits.

The column vector $\tilde{\mathbf{h}}$ contains channel coefficients in the frequency domain corresponding to all the N_t transmit antennas. The entries of $\tilde{\mathbf{h}}$ are modeled as complex Gaussian random variables with zero mean and unit variance. Specifically, let $\tilde{\mathbf{h}}_i = \left[\tilde{h}_i(1), \tilde{h}_i(2), \cdots, \tilde{h}_i(F)\right]^T$ denote the channel coefficients for the ith transmit antenna. Then $\tilde{\mathbf{h}} = \left[\tilde{\mathbf{h}}_1^T, \tilde{\mathbf{h}}_2^T, \cdots, \tilde{\mathbf{h}}_{N_t}^T\right]^T$.

As mentioned, the position of the sub-matrix $\mathbf{X}_q$ in $\mathbf{X}$ is determined by p_1 bits (i.e., the antenna that is activated for the group). Given the index of the active antenna q, the receive signal $\mathbf{y}$ in (4.1) can be rewritten as

$$\mathbf{y} = \mathbf{X}_q\tilde{\mathbf{h}}_q + \mathbf{w}. \tag{4.3}$$

Let L be the length of the multipath channel in the time domain and define $\bar{\mathbf{h}}_i = [\bar{h}_i(1), \ldots, \bar{h}_i(L)]^T$ to be the time-domain channel impulse response. Then the channel vector $\tilde{\mathbf{h}}_i$ in the frequency domain and $\bar{\mathbf{h}}_i$ in the time domain is related by $\tilde{\mathbf{h}}_i = \mathbf{Q}\bar{\mathbf{h}}_i$, where $\mathbf{Q}$ is the $F \times L$ truncated FFT (Vandermonde) matrix. That is, $\mathbf{Q} = [\mathbf{q}(1), \cdots, \mathbf{q}(F)]^T$ with $\mathbf{q}(k) = [1, \exp(-j2\pi k/F), \cdots, \exp(-j2\pi k(L-1)/F)]$. Assume that the channel vector $\bar{\mathbf{h}}_i$ is zero mean, complex Gaussian with *full rank* correlation matrix $\mathbf{R}_h \triangleq E(\bar{\mathbf{h}}_i \bar{\mathbf{h}}_i^H)$. Since $\mathbf{R}_h$ is positive definite Hermitian symmetric, it can be expressed as $\mathbf{R}_h = \mathbf{B}\mathbf{B}^H$. Then $\bar{\mathbf{h}}_i$ can be expressed as $\bar{\mathbf{h}}_i = \mathbf{B}\mathbf{h}_i$ with the entries of $\mathbf{h}_i$ being i.i.d zero mean complex Gaussian random variables with unit variance. Since $\tilde{\mathbf{h}}_i = \mathbf{Q}\mathbf{B}\mathbf{h}_i$, the relationship between $\mathbf{h} = \left[\mathbf{h}_1^T, \mathbf{h}_2^T, \cdots, \mathbf{h}_{N_t}^T\right]^T$ in the time domain and $\tilde{\mathbf{h}} = \left[\tilde{\mathbf{h}}_1^T, \tilde{\mathbf{h}}_2^T, \cdots, \tilde{\mathbf{h}}_{N_t}^T\right]^T$ in the frequency domain is as follows:

$$\tilde{\mathbf{h}} = \mathbf{V}\mathbf{h}, \text{ where } \mathbf{V} = \begin{pmatrix} \mathbf{QB} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{QB} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{QB} \end{pmatrix} \tag{4.4}$$

Given the signal model in (4.1) and (4.4), the ML detection of the transmitted signal matrix $\mathbf{X}$ is to solve the following constrained least-square problem:

$$\tilde{\mathbf{X}}_{\mathrm{ML}} = \arg\min_{\mathbf{X}} \left\| \mathbf{y} - \mathbf{X}\tilde{\mathbf{h}} \right\|^2. \tag{4.5}$$

4.1.2 Performance Analysis

This section examines the performance of the proposed SM-OFDM-GLCP system under the ML decoding rule in (4.5), where the emphasis is on the diversity and coding gains. To this end, the conditional PEP is given as

$$P\left(\mathbf{X} \to \hat{\mathbf{X}} \mid \mathbf{h}\right) = Q\left(\sqrt{\frac{\delta_e}{2N_0}}\right) \tag{4.6}$$

where $\delta_e = \|(\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}}\|_F^2 = \tilde{\mathbf{h}}^H\mathbf{A}_e\tilde{\mathbf{h}} = \mathbf{h}^H\mathbf{V}^H\mathbf{A}_e\mathbf{V}\mathbf{h}$ and $\mathbf{A}_e = (\mathbf{X}-\hat{\mathbf{X}})^H(\mathbf{X}-\hat{\mathbf{X}})$. The derivation for (4.6) is as follows:

$$
\begin{aligned}
& P\left(\mathbf{X}\to\hat{\mathbf{X}} \mid \tilde{\mathbf{h}}\right) \\
&= P\left\{\| \mathbf{y}-\mathbf{X}\tilde{\mathbf{h}} \|^2 - \| \mathbf{y}-\hat{\mathbf{X}}\tilde{\mathbf{h}} \|^2 > 0 \mid \tilde{\mathbf{h}}\right\} \\
&= P\left\{(\mathbf{y}-\mathbf{X}\tilde{\mathbf{h}})^H(\mathbf{y}-\mathbf{X}\tilde{\mathbf{h}}) - (\mathbf{y}-\hat{\mathbf{X}}\tilde{\mathbf{h}})^H(\mathbf{y}-\hat{\mathbf{X}}\tilde{\mathbf{h}}) > 0 \mid \tilde{\mathbf{h}}\right\} \\
&= P\left\{[(\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}}+\mathbf{w}]^H[(\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}}+\mathbf{w}] - \mathbf{w}^H\mathbf{w} < 0 \mid \tilde{\mathbf{h}}\right\} \\
&= P\Bigg\{\tilde{\mathbf{h}}^H(\mathbf{X}-\hat{\mathbf{X}})^H(\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}} \\
&\qquad - \underbrace{[\mathbf{w}^H(\hat{\mathbf{X}}-\mathbf{X})\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H(\hat{\mathbf{X}}-\mathbf{X})^H\mathbf{w}]}_{T} < 0 \mid \tilde{\mathbf{h}}\Bigg\} \\
&= P\left\{T > \| (\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}} \|^2 \mid \tilde{\mathbf{h}}\right\} \\
&= Q\left(\frac{\| (\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}} \|^2}{\sqrt{2N_0 \| (\mathbf{X}-\hat{\mathbf{X}})\tilde{\mathbf{h}} \|^2}}\right) = Q\left(\sqrt{\frac{\delta_e}{2N_0}}\right).
\end{aligned}
\tag{4.7}
$$

The function $Q(x)$ can be approximated as [35]:

$$
Q(x) \simeq \frac{1}{12}\mathrm{e}^{-x^2/2} + \frac{1}{4}\mathrm{e}^{-2x^2/3} \tag{4.8}
$$

Hence, the unconditional PEP (UPEP) can be evaluated as:

$$
P\left(\mathbf{X}\to\hat{\mathbf{X}}\right) \simeq E_{\mathbf{h}}\left\{\frac{1}{12}\mathrm{e}^{-q_1\mathbf{h}^H\mathbf{V}^H\mathbf{A}_e\mathbf{V}\mathbf{h}} + \frac{1}{4}\mathrm{e}^{-q_2\mathbf{h}^H\mathbf{V}^H\mathbf{A}_e\mathbf{V}\mathbf{h}}\right\} \tag{4.9}
$$

where $q_1 = 1/4N_0$ and $q_2 = 1/3N_0$.

Considering the case that the multipath channel coefficients in the time domain are uncorrelated, i.e., $E\{\mathbf{h}\mathbf{h}^H\} = \mathbf{I}$, then the pdf of $\mathbf{h}$ is $f(\mathbf{h}) = \pi^{-LN_t}\mathrm{e}^{-\mathbf{h}^H\mathbf{h}}$. Thus, the UPEP

is

$$P\left(\mathbf{X}\to\hat{\mathbf{X}}\right)\simeq\int_{\mathbf{h}}\frac{1}{12}\mathrm{e}^{-q_1\mathbf{h}^H\mathbf{V}^H\mathbf{A}_e\mathbf{V}\mathbf{h}}f(\mathbf{h})d\mathbf{h}$$
$$+\int_{\mathbf{h}}\frac{1}{4}\mathrm{e}^{-q_2\mathbf{h}^H\mathbf{V}^H\mathbf{A}_e\mathbf{V}\mathbf{h}}f(\mathbf{h})d\mathbf{h} \tag{4.10}$$
$$=\frac{\pi^{-LN_t}}{12}\int_{\mathbf{h}}\mathrm{e}^{-\mathbf{h}^H[\mathbf{I}+q_1\mathbf{V}^H\mathbf{A}_e\mathbf{V}]\mathbf{h}}d\mathbf{h}$$
$$+\frac{\pi^{-LN_t}}{4}\int_{\mathbf{h}}\mathrm{e}^{-\mathbf{h}^H[\mathbf{I}+q_2\mathbf{V}^H\mathbf{A}_e\mathbf{V}]\mathbf{h}}d\mathbf{h} \tag{4.11}$$
$$=\frac{1/12}{\det(\mathbf{I}+q_1\mathbf{V}^H\mathbf{A}_e\mathbf{V})}+\frac{1/4}{\det(\mathbf{I}+q_2\mathbf{V}^H\mathbf{A}_e\mathbf{V})} \tag{4.12}$$
$$=\frac{1/12}{\prod_{i=1}^{r}(1+q_1\lambda_i(\mathbf{C}_e))}+\frac{1/4}{\prod_{i=1}^{r}(1+q_2\lambda_i(\mathbf{C}_e))} \tag{4.13}$$

where $\mathbf{C}_e=\mathbf{V}^H\mathbf{A}_e\mathbf{V}$, $r=\mathrm{rank}(\mathbf{C}_e)$ and $\lambda_i(\mathbf{C}_e)$ is the ith eigenvalue of $\mathbf{C}_e$.

For high SNR values, the UPEP can be approximated as

$$P\left(\mathbf{X}\to\hat{\mathbf{X}}\right)\simeq\left(12q_1^r\prod_{i=1}^{r}\lambda_i(\mathbf{C}_e)\right)^{-1}+\left(4q_2^r\prod_{i=1}^{r}\lambda_i(\mathbf{C}_e)\right)^{-1} \tag{4.14}$$

Once the UPEP is evaluated, the average bit error probability (ABEP) can be calculated as

$$P_b\simeq\frac{1}{(p_1+p_2)n_{\mathbf{X}}}\sum_{\mathbf{X}}\sum_{\hat{\mathbf{X}}}P\left(\mathbf{X}\to\hat{\mathbf{X}}\right)\chi\left(\mathbf{X},\hat{\mathbf{X}}\right), \tag{4.15}$$

where $n_{\mathbf{X}}$ is the number of possible realizations of $\mathbf{X}$ and $\chi(\mathbf{X},\hat{\mathbf{X}})$ represents the number of bit errors for the corresponding pairwise error event.

The rank r of matrix $\mathbf{C}_e$ is known as the *diversity order*, denoted as $G_{d,e}$, while $\prod_{i=1}^{r}\lambda_i(\mathbf{C}_e)$ is called the *coding advantage*, denoted as $G_{c,e}$. It is noted that both $G_{d,e}$ and $G_{c,e}$ depend on the choice of pair $(\mathbf{X},\hat{\mathbf{X}})$. More generally, the *diversity order* and *coding gain* of the system are defined as

$$G_d=\min_{\forall\mathbf{X}\neq\hat{\mathbf{X}}}G_{d,e}=\min_{\forall\mathbf{X}\neq\hat{\mathbf{X}}}\mathrm{rank}(\mathbf{C}_e),\ \text{and}\ G_c=\min_{\forall\mathbf{X}\neq\hat{\mathbf{X}}}G_{c,e} \tag{4.16}$$

Following the same analysis as in [53], it can be shown that under the necessary condition that $F\geq L$, the use of a properly-designed rotation matrix Θ always guarantee that the

minimum rank of matrix $\mathbf{V}^H\mathbf{A}_e\mathbf{V}$ is equal to L, for all $\mathbf{A}_e$. Thus, by choosing the group size properly and performing appropriate rotation, the diversity gain of SM-OFDM-GLCP is always L, which is the same as that of OFDM-GLCP. However, by exploiting spatial modulation, SM-OFDM-GLCP enhances the transmission rate because more information bits are conveyed by antenna indexes. Specifically, the number of information bits per subcarrier is given as $\log_2 M + \frac{\lfloor \log_2 N_t \rfloor}{F}$, where M is the constellation size.

4.1.3 Simulation Results

In this section, performance comparisons among SM-OFDM-GLCP, the conventional SM-OFDM [13], and OFDM-GLCP are performed. In all simulations, the maximum likelihood detection is implemented. The number of channel taps is $L = 3$, while the total number of subcarriers is set to $N = 16$ and one receive antenna is employed. All 16 subcarriers are divided into 4 groups with $F = 4$ subcarriers per group (hence satisfying the condition $F \geq L$). It is noted that practical OFDM systems employ larger numbers of subcarriers. However, for $N > 16$ the results would be the same as long as $F = 4$. This is because the performance of the whole system is the same as the performance of each subcarrier group.

To simulate a correlated fading channel, the following correlation matrix is assumed:

$$\mathbf{R}_h = \frac{1}{L}\begin{pmatrix} 1 & 0.6 & 0 \\ 0.6 & 1 & 0.6 \\ 0 & 0.6 & 1 \end{pmatrix} \tag{4.17}$$

All simulation results in this section are presented in terms of the bit error rate (BER) versus E_b/N_0, where E_b is the transmitted energy per information bit.

In Figure 4.3, both SM-OFDM-GLCP and OFDM-GLCP use symbols drawn from QPSK but they employ different numbers of transmit antennas to achieve 2.5 bits/subcarrier and 2 bits/subcarrier, respectively. Both schemes utilize the linear precoder to realize the maximum multipath diversity gain of $L = 3$. As can be seen from the figure, SM-OFDM-GLCP yields a better performance at high E_b/N_0. The reason for this is that information bits are more effectively transmitted in both spatial and signal space dimensions in SM-OFDM-GLCP as opposed to only the signal space dimension in OFDM-GLCP. The figure also plots

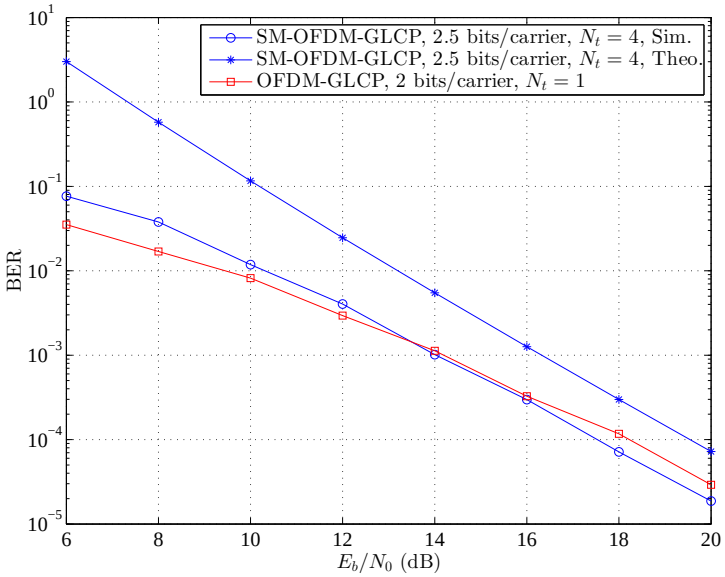

Figure 4.3 BER performance comparison between SM-OFDM-GLCP and OFDM-GLCP.

the theoretical upper (union) bound of SM-OFDM-GLCP, which appears quite loose. Nevertheless it can be seen that the slope of the upper bound curve is asymptotically the same as the slope of the corresponding simulation curve, which confirms the diversity analysis presented before.

To obtain the BER curves presented in Figure 4.4, for 4 transmit antennas, the information symbols in the conventional SM-OFDM are drawn from QPSK on each subcarrier, the information symbols are 16-QAM in the proposed SM-OFDM-GLCP. The main difference between SM-OFDM-GLCP and SM-OFDM is the number of bits conveyed on each subcarrier by antenna indexing. While SM-OFDM with spatial modulation applied on each subcarrier conveys 2 bits/subcarrier by antenna indexing (4 transmit antennas), SM-OFDM-GLCP only conveys 2 additional bits per subcarrier group. Since the group size is $F = 4$, SM-OFDM-GLCP conveys additional 0.5 bit/subcarrier by antenna indexing. However, this disadvantage is compensated by the ability to employ a higher-order QAM constellation, thanks to the transmit diversity gain realized with a well-designed rotation matrix Θ. Due to the higher diversity gain, it is clear from Figure 4.4 that SM-OFDM-GLCP yields much better performance than the conventional SM-OFDM.

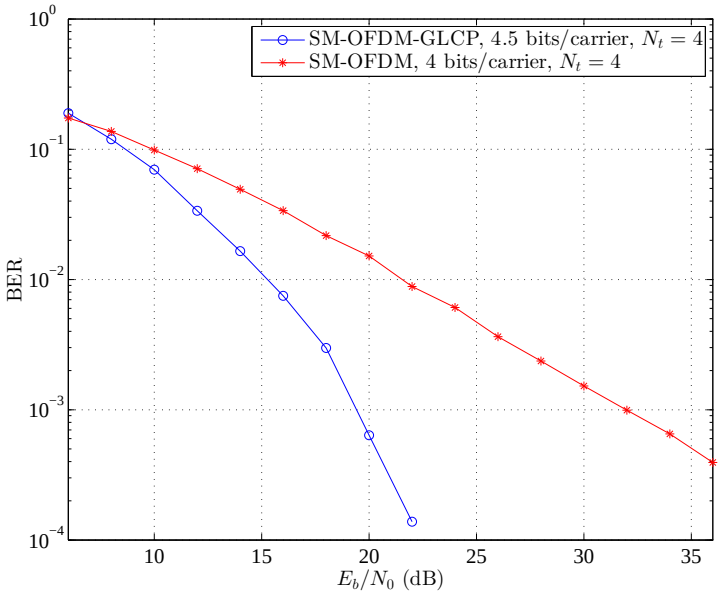

Figure 4.4 BER performance comparison between SM-OFDM-GLCP and SM-OFDM.

4.2 Performance Comparison of IM-Based OFDM and Precoded-OFDM

Recall that, in spatial modulation, only a subset of available transmit antennas is activated to transmit constellation symbols, while antenna indices are used to send extra information bits. Motivated by the idea of spatial modulation, various modifications were proposed for OFDM in which, instead of antenna indices, subcarrier indices are employed to send extra information bits. Specifically, in OFDM with index modulation (OFDM-IM) [35], subcarriers are classified to be active or inactive. The choice of active subcarriers is determined by a group input bits. The active subcarriers transmit constellation symbols while inactive subcarriers transmit nothing (i.e., being turned off). More recently, a dual-mode OFDM-IM scheme is proposed as an extension of OFDM-IM [57, 58]. The extension is that a different constellation is used for subcarriers that are supposed to be inactive in OFDM-IM. In doing so, the overall spectral efficiency is improved. It is noted that there are still indexing bits in dual-mode OFDM-IM and for the scheme to work well it is important to design proper constellations for the two modes of subcarriers.

The reason that the IM-based OFDM systems have a better performance than the classical OFDM system is because a higher diversity gain is achieved for bits transmitted by means of subcarrier indexing. The another advantage of the IM-based OFDM systems is that they are more flexible than the classical OFDM system in terms of spectral efficiency. This is because in a classical OFDM system the modulated symbols are usually drawn from a unique signal constellation, whereas in IM-based OFDM systems the number of active subcarriers can be chosen quite arbitrarily.

It is emphasized that the performance advantages of the IM-based OFDM systems are demonstrated in [35] and [57] against the classical OFDM. As mentioned before, precoded OFDM is well known to have a better performance than the classical OFDM due to its ability to gather the frequency diversity through subcarrier grouping and spreading. On the other hand, to yield more flexible spectral efficiency, one can employ different constellations for different subcarriers. As such, the main objective of this section is to compare IM-based OFDM with precoded-OFDM with multiple constellations (precoded-OFDM-MConst) so that a better understanding of the advantages and disadvantages of IM-based OFDM systems can be established.

4.2.1 System Model

At the transmitter of precoded-OFDM-MConst, a total number of B bits are packed in each OFDM symbol. These B bits are divided into G groups, each group having $p = B/G$ bits. Each group is assigned to an OFDM sub-symbol with length F. Hence, the total number of subcarriers in an OFDM symbol (or the length of the FFT) is $N = FG$. In precoded-OFDM-MConst, all subcarriers are active and there are no subcarrier indexing bits. The p bits associated with the gth OFDM sub-symbol go through the M-ary mapper to be mapped to F signal constellation symbols. The output of the mapper is represented as $\mathbf{s}^{(g)} = \{s_1^{(g)}, \cdots, s_F^{(g)}\}$, where $g \in \{1, \cdots, G\}$ and $s_i^{(g)} \in S$ with S is the set of constellation symbols. These F symbols are then fed to a precoder $\mathbf{\Theta}$, which is a $F \times F$ matrix, to maximize the diversity gain. The output of the precoder is $\mathbf{X}^{(g)} = [X_1^{(g)}, X_2^{(g)}, \cdots, X_F^{(g)}]^T = \mathbf{\Theta}\mathbf{s}^{(g)}$. In this thesis, the precoder $\mathbf{\Theta}$ is designed by the LCP-A method proposed in [53], which can

be written as a Vandermonde matrix as

$$\mathbf{\Theta} = \frac{1}{\beta}\begin{pmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{F-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{F-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_F & \cdots & \alpha_F^{F-1} \end{pmatrix} \tag{4.18}$$

where β is a normalization factor chosen to satisfy the power constraint $\mathrm{tr}(\mathbf{\Theta}\mathbf{\Theta}^H) = F$. For example, with $F = 2$ the design is $\alpha_1 = \mathrm{e}^{-j\frac{\pi}{4}}$ and $\alpha_2 = \mathrm{e}^{-j\frac{5\pi}{4}}$. On the other hand, with $F = 4$, one has $\alpha_1 = \mathrm{e}^{-j\frac{\pi}{8}}$, $\alpha_2 = \mathrm{e}^{-j\frac{5\pi}{8}}$, $\alpha_3 = \mathrm{e}^{-j\frac{9\pi}{8}}$, $\alpha_4 = \mathrm{e}^{-j\frac{13\pi}{8}}$.

After precoding, the G precoded OFDM sub-symbols are arranged into a length-N vector $\mathbf{X} = [X_1, X_2, \cdots, X_N]^T$ according to the interleaved subcarrier grouping scheme [53, 56]. The vector $\mathbf{X}$ is then processed by IFFT to generate the time-domain symbol vector $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T = \mathrm{IFFT}(\mathbf{X})$. Then a CP with length L is added before the signal samples are fed into a parallel-to-serial converter and a digital-to-analog converter to generate an OFDM symbol for transmission. Let the channel impulse response (CIR) coefficients of the frequency-selective Rayleigh fading channel be represented as $\mathbf{h} = [h_1, h_2, \cdots, h_V]^T$, where $h_i \sim \mathcal{CN}(0, \frac{1}{V})$ for $i = 1, \cdots, V$ and V is the number of channel taps ($V \leq L$). The frequency-domain channel coefficients are defined as the N-point FFT of $\bar{\mathbf{h}}$, given as $\mathbf{H} = [H_1, H_2, \cdots, H_N]^T = \frac{1}{\sqrt{N}}\mathrm{FFT}(\bar{\mathbf{h}})$, where $\bar{\mathbf{h}} = [h_1, h_2, \cdots, h_V, 0, \cdots, 0]^T$ is the length-N vector. Assume further that the channel remains constant during the transmission of an OFDM symbol. Then it is well known that output of the FFT at the receiver, Y_i, is related to the ith element of vector $\mathbf{X}$ as

$$Y_i = H_i X_i + W_i, \qquad i = 1, \cdots, N \tag{4.19}$$

where $W_i \sim \mathcal{CN}(0, N_{0,F})$ is the frequency-domain channel noise sample and $N_{0,F}$ is the noise variance. Furthermore, the received signal vector for the gth OFDM sub-symbol can be expressed as:

$$\mathbf{Y}^{(g)} = \mathrm{diag}\{\mathbf{H}^{(g)}\}\mathbf{X}^{(g)} + \mathbf{W}^{(g)} = \mathrm{diag}\{\mathbf{H}^{(g)}\}\mathbf{\Theta}\mathbf{s}^{(g)} + \mathbf{W}^{(g)} \tag{4.20}$$

where $\mathbf{H}^{(g)} = [H_1^{(g)}, H_2^{(g)}, \cdots, H_F^{(g)}]^T$ collects the frequency-domain channel coefficients corresponding to the gth group, and $\mathbf{W}^{(g)} = [W_1^{(g)}, W_2^{(g)}, \cdots, W_F^{(g)}]^T$. The maximum likelihood

(ML) detection of the information symbols of the gth OFDM sub-symbol is as follows:

$$\mathbf{s}_{\mathrm{ML}}^{(g)} = \arg \quad \min_{\forall \mathbf{s}^{(g)}} \|\mathbf{Y}^{(g)} - \mathrm{diag}\{\mathbf{H}^{(g)}\}\boldsymbol{\Theta}\mathbf{s}^{(g)}\| \tag{4.21}$$

It is pointed out that the complexity of the ML detection depends on the group size F, as well as the size of the constellations used in each group.

4.2.2 Simulation Results and Comparisons

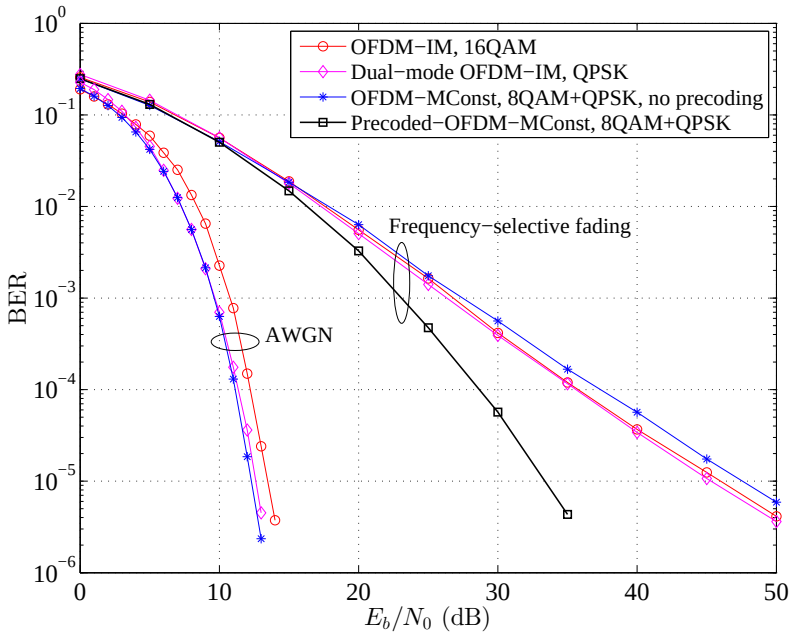

Figure 4.5 BER performance comparison of the precoded-OFDM-MConst, OFDM-IM and dual-mode OFDM-IM at the spectral efficiency of 10 bits/group.

In this section, the simulation results are presented to compare the performance of precoded-OFDM-MConst and the two IM-based OFDM systems, namely OFDM-IM and dual-mode OFDM-IM. The frequency-selective Rayleigh fading channel used in the simulation is similar to that in [35,57] and has a CIR length of $V = 10$. The number of subcarriers is set to $N = 128$, which is divided into $G = 32$ groups with $F = 4$ subcarriers per group. The cyclic prefix length is chosen to be $L = 16$ (again, similar to [35, 57]). All systems are simulated and compared under two channel scenarios: one with additive white Gaussian noise (AWGN) only, and one under the presence of frequency-selective Rayleigh fading.

The signal-to-noise ratio (SNR) is defined as E_b/N_0, where E_b is the transmitted energy per information bit and N_0 is one-sided power spectral density of AWGN. Over an AWGN channel, the relative error performance of different systems can be judged solely based on the minimum Euclidean distance normalized by the transmitted bit energy, defined as:

$$d_{\min} = \min_{i,j} \sqrt{\frac{1}{E_b}} \|\mathbf{x}_i - \mathbf{x}_j\| \tag{4.22}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are two different realizations of an OFDM sub-symbol. On the other hand, over a frequency-selective fading channel, the diversity gain plays a more important role than $d_{\min}$ in the high SNR region.

In Figure 4.5, precoded-OFDM-MConst is compared with OFDM-IM and dual-mode OFDM-IM at the spectral efficiency of 10 bits/group. To achieve such a spectral efficiency, OFDM-IM activates 2 out of 4 subcarriers (i.e., 2 subcarriers are inactive). Each active subcarrier in OFDM-IM sends a 16QAM constellation symbol. For the dual-mode OFDM-IM, 2 out of 4 subcarriers in each group send symbols drawn from the first QPSK constellation while the remaining 2 subcarriers send symbols drawn from the second QPSK constellation. Two QPSK signal constellations used in dual-mode OFDM-IM are optimized such that the minimum Euclidean distance between any two OFDM symbols is maximized. For the precoded-OFDM-MConst, in each group, 2 subcarriers send 2 symbols taken from 8QAM constellation and the remaining 2 subcarriers sends 2 symbols taken from a QPSK constellation. To enlarge the minimum Euclidean distance between any two OFDM syb-symbols corresponding to the same set of subcarriers, the constellation points of 8QAM are $\{2, 2+2j, 2j, -2+2j, -2, -2j, 2-2j\}$, while those of QPSK are $\{1+j, -1+j, -1, -j, 1-j\}$. For precoded-OFDM-MConst it is found that $d_{\min} = 1.4907$, while $d_{\min}$ of OFDM-IM is 1.3333 and that of dual-mode OFDM-IM is 1.3706. As such, it is expected that the precoded-OFDM-MConst has the best performance, followed by OFDM-IM and dual-mode OFDM-IM. This expectation is confirmed by the simulation results presented in Figure 4.5.

For the case of a frequency-selective Rayleigh fading channel, the dual-mode OFDM-IM system performs better than the OFDM-IM system, the same observatio made in [57]. Both the dual-mode OFDM-IM and OFDM-IM systems use subcarrier indices to transmit information bits in addition to those information bits transmitted by constellation symbols.

However, the dual-mode OFDM-IM system utilizes "second-mode" subcarriers to send more information symbols that are drawn from a second constellation. This make $d_{\min}$ of the dual-mode OFDM-IM larger than $d_{\min}$ of OFDM-IM. Performance of OFDM using multiple constellations is studied for two cases: one with LCP-A precoding and one without precoding. Without precoding, the performance of OFDM-MConst is worse than that of OFDM-IM and dual-mode OFDM-IM. This is because both OFDM-IM and dual-mode OFDM-IM have a higher diversity gain for the subcarrier indexing bits [35], which becomes more important than the larger $d_{\min}$ of OFDM-MConst in a frequency-selective fading channel. For precoded-OFDM-MConst using LCP-A precoder, its performance is significantly better than that of OFDM-IM and dual-mode OFDM-IM. This is because the precoded-OFDM-MConst system achieves the diversity order of four which is equal to the length of OFDM sub-symbol $F = 4$. Such a high diversity order is enjoyed by all the information bits, not only those sent by subcarrier indices as in the two IM-based OFDM schemes. At the BER level of 10^{-5} the SNR gain (i.e., coding gain) of the precoded-OFDM-MConst over the other two IM-based OFDM schemes is about 11 dB.

Similar performance comparison among three systems is presented in Figure 4.6, but at a higher spectral efficiency of 18 bits/group. To pack 18 bits in each subcarrier group, OFDM-IM activates 2 out of 4 subcarriers in one group and each subcarrier transmits a 256QAM constellation symbol. For the dual-mode OFDM-IM system, in one group 2 subcarriers are chosen to transmit symbols taken from the first 16QAM constellation, while the other 2 subcarriers transmit symbols taken from the second 16QAM constellation. For precoded-OFDM-MConst, 4 subcarriers in one group transmit symbols taken from 32QAM and 16QAM. Specifically, the first two subcarriers transmits 32QAM symbols and the last two subcarriers transmits 16QAM symbols. The minimum distances are found to be 0.9428, 0.8944 and 0.4339 for the precoded-OFDM-MConst, dual-mode OFDM-IM and OFDM-IM, respectively. These minimum distance values explain why precoded-OFDM-MConst and dual-mode OFDM-IM perform similarly and much better than OFDM-IM over an AWGN channel. More importantly, under the frequency selective Rayleigh fading channel, the precoded-OFDM-MConst significantly outperforms the two IM-based OFDM schemes,

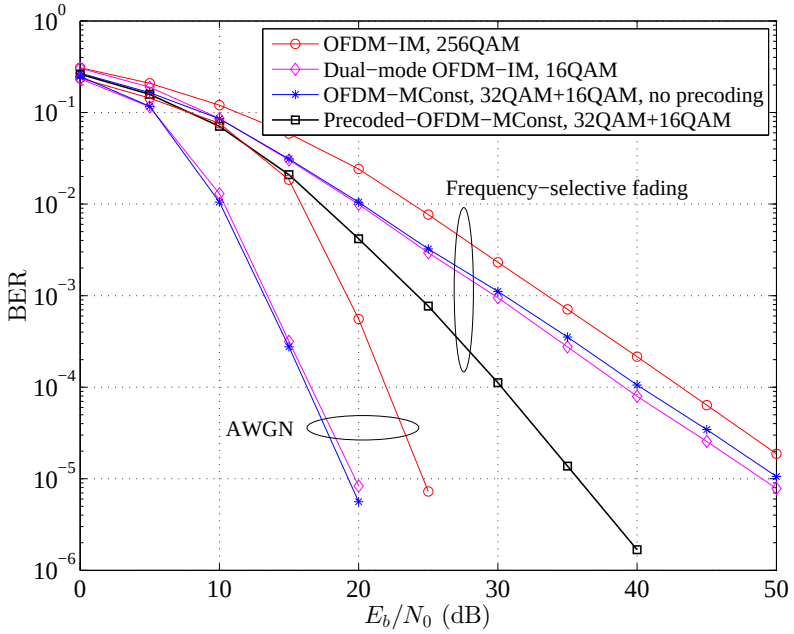

Figure 4.6 BER performance comparison of the precoded-OFDM-MConst, OFDM-IM and dual-mode OFDM-IM at the spectral efficiency of 18 bits/group.

thanks to its diversity order of 4 that is enjoyed by all information bits. In particular, at the BER level of 10^{-5} the coding gain of the precoded-OFDM-MConst over the other two IM-based OFDM schemes is about 14 dB in this case.

4.3 Summary

In this chapter, the SM-OFDM-GLCP scheme was proposed so that SM can be used effectively in OFDM systems that employ multiple transmit antennas and operate over frequency-selective Rayleigh fading channels. It is shown that both diversity and coding gain are achieved by the proposed scheme which is not the case for the conventional SM-OFDM. Next, OFDM with multiple constellations and IM-based OFDM were studied. The results show that the precoded-OFDM with multiple constellations significantly outperforms the precoded IM-based OFDM schemes. This finding suggests that one needs to carefully consider the performance complexity tradeoff when choosing between IM-based OFDM and OFDM with multiple constellations.

5. Multiuser Spatial Modulation

Many SM-MIMO schemes have been studied by implicitly assuming a single-user transmission. However, this operating scenario is quite restrictive for typical cellular deployments, where many users may simultaneously transmit over the same resource block, aiming at maximizing the aggregate throughput at the cost of increasing the interference. Motivated by this consideration, the focus of this chapter is to investigate SM for multiuser communications.

In the context of multiuser communications, the performance of both optimal and suboptimal receivers designed for SM-MIMO communications has been investigated in the presence of multiple-access interference [59]. The authors in [60] study the error probability of SSK-MIMO by considering two receivers: (1) the single-user receiver, which is of low complexity, but it is oblivious of the interference; and (2) the optimum ML multiuser receiver, which is of high complexity, and has the benefit of being interference aware. The authors show that the error floor of the single-user receiver can be significantly reduced by increasing the number of receive antennas (RAs). In particular, if the number of RAs goes to infinity, the error probability goes asymptotically to zero. This behavior is known as "massive" MIMO effect [61]. A layered SM scheme is proposed in [62] as a low-complexity scheme for multiuser downlink systems where the spatial domain is used to activate TAs to transmit information to several users simultaneously. In [63], the precoding-aided SM is applied to multiuser downlink transmission to resist a multiple-antenna eavesdropper. In particular, the authors design the signal precoding matrices to cancel the multiuser interference and modulate partial information bits on the indices of RAs. For uplink transmission, the achievable uplink spectral efficiency of a multicell massive SM-MIMO system relying on linear combining schemes is investigated in [64]. In [65], a low-complexity message passing de-quantization

detector for a massive SM-MIMO system with low-resolution analog-to-digital converters is proposed for multiuser detection. In [66], a detector based on the compressive sensing principle is presented to improve the performance of the conventional detection algorithm by exploiting the structure and sparsity of the SM transmitted signals in the multiple-access channels.

This chapter is concerned with downlink multiuser transmission between a base station and multiple users. Motivated by [63] in which downlink transmission implements precoding-aided SM to mitigate the multiuser interference, a novel scheme, called multiuser precoding-aided quadrature space shift keying (MU-PQSSK) is developed. Although the modulation scheme used for all users is SSK, the even-indexed and odd-indexed users are distinguished by means of inphase and quadrature SSK, respectively. In fact, the advantage of combining inphase and quadrature SSK have recently been investigated in different systems [48,67] and this chapter is another investigation of such advantage in the context of multiuser downlink transmission. Here, the advantage is to reduce the minimum number of TAs required at base station in such a way that a precoder matrix can still be used at the base station to mitigate multiuser interference. In addition to the maximum likelihood detection, the low-complexity zero forcing (ZF) detection method is also investigated for the proposed MU-PQSSK system.

The remaining of this chapter is organized as follows. Section 5.1 reviews the downlink multiuser transmission system using the precoding-aided SSK which is modified from [63]. Section 5.2 presents the proposed MU-PQSSK together with the ML and ZF receivers. Section 5.3 provides performance analysis. Simulation results and performance comparisons are discussed in Section 5.4.

Notation: Bold letters are used for column vectors, while capital bold letters are for matrices. The operators $(\cdot)^T$ and $(\cdot)^H$ denote transposition and Hermitian transposition, respectively. $\|\cdot\|$ stands for the Frobenius norm. $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ the are real and imaginary parts of a complex vector. $\det(\cdot)$ and $\mathrm{Tr}(\cdot)$ are determinant and trace of a square matrix.

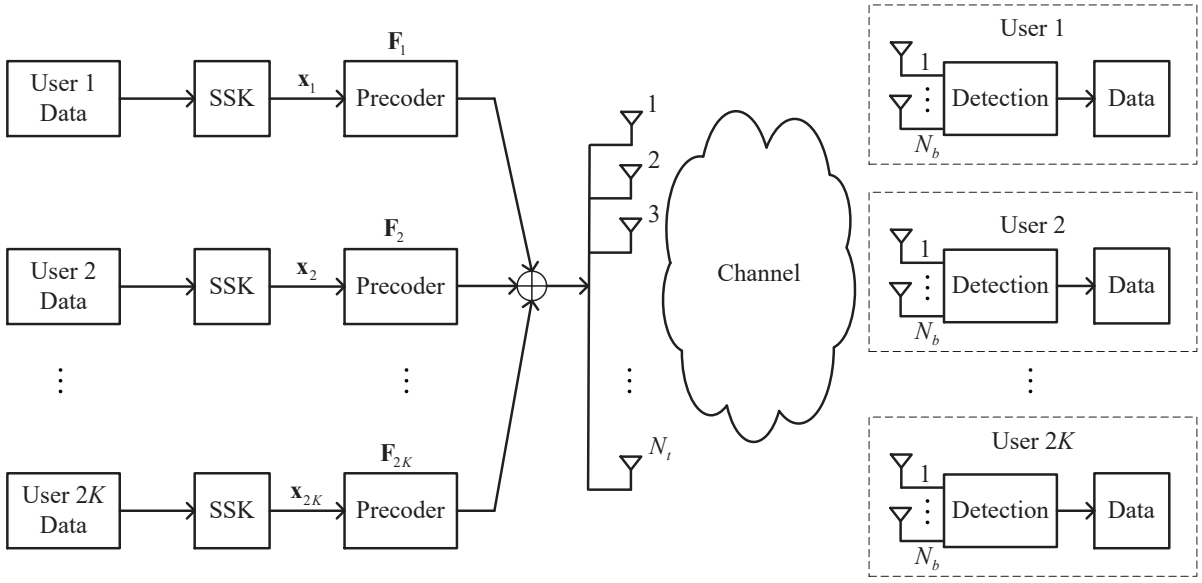

Figure 5.1 Transceiver of a MU-PSSK system.

5.1 Overview of MU-PSSK

The multiuser downlink transmission system using SSK with precoder is called MU-PSSK. It is basically the same as the system discussed in [63] but with SM replaced by SSK. As illustrated in Figure 5.1, the base station is equipped with N_t TAs to send data to $2K$ users and each user has N_b RAs. At the base station, the information bits of user k go through the SSK mapper to obtain an $N_a \times 1$ output vector $\mathbf{x}_k$:

$$\mathbf{x}_k = [0 \cdots 0 \quad 1 \quad 0 \cdots 0]^T \in \mathbb{C}^{N_a \times 1} \tag{5.1}$$

The super symbol containing information for all users at the base station can be expressed as $\mathbf{x} = [\mathbf{x}_1^T \cdots \mathbf{x}_k^T \cdots \mathbf{x}_{2K}^T]^T \in \mathbb{C}^{2KN_a \times 1}$. Let $\mathbf{H}$ be the over all channel matrix, which can be written as

$$\mathbf{H} = [\mathbf{H}_1^T \cdots \mathbf{H}_k^T \cdots \mathbf{H}_{2K}^T]^T \in \mathbb{C}^{N_r \times N_t} \tag{5.2}$$

where $N_r = 2KN_b$ is the total number of RAs of all users and $\mathbf{H}_k \in \mathbb{C}^{N_b \times N_t}$ is the channel matrix between the base station and user k, whose entries are independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and unit variance.

To mitigate the interference from different users, at the base station user k employs a precoder matrix $\mathbf{F}_k \in \mathbb{C}^{N_t \times N_a}$. Define $\mathbf{F} = [\mathbf{F}_1 \cdots \mathbf{F}_k \cdots \mathbf{F}_{2K}]$. Then the received signal

vector for all users can be expressed as

$$\mathbf{y} = \mathbf{HFx} + \mathbf{n}, \tag{5.3}$$

where $\mathbf{y} = [\mathbf{y}_1^T \cdots \mathbf{y}_k^T \cdots \mathbf{y}_{2K}^T]^T \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{y}_k \in \mathbb{C}^{N_b \times 1}$ is the received signal vector at user k, $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is the noise vector whose entries are i.i.d complex Gaussian random variables with zero mean and variance N_0. The precoder matrix $\mathbf{F}$ is designed such that [63]

$$\mathbf{HF} = \begin{bmatrix} \mathbf{H}_1\mathbf{F}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2\mathbf{F}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_{2K}\mathbf{F}_{2K} \end{bmatrix}. \tag{5.4}$$

Equivalently, one needs to design $\mathbf{F}_k$ such that $\tilde{\mathbf{H}}_k\mathbf{F}_k = \mathbf{0}$ where

$$\tilde{\mathbf{H}}_k = [\mathbf{H}_1^T \cdots \mathbf{H}_{k-1}^T \mathbf{H}_{k+1}^T \cdots \mathbf{H}_{2K}^T]^T \in \mathbb{C}^{(N_r - N_b) \times N_t} \tag{5.5}$$

If $\mathbf{F}_k$ belongs to the null space of $\tilde{\mathbf{H}}_k$, then $\tilde{\mathbf{H}}_k\mathbf{F}_k = \mathbf{0}$. To find the null space of $\tilde{\mathbf{H}}_k$, using singular value decomposition (SVD), $\tilde{\mathbf{H}}_k$ can be expressed as

$$\tilde{\mathbf{H}}_k = \mathbf{UE}[\mathbf{V}^{(1)}\mathbf{V}^{(0)}]^H \tag{5.6}$$

where $\mathbf{U} \in \mathbb{C}^{(N_r - N_b) \times (N_r - N_b)}$ is the left singular matrix of $\tilde{\mathbf{H}}_k$, $\mathbf{E} \in \mathbb{C}^{(N_r - N_b) \times N_t}$ is a rectangular matrix with its diagonal consisting of the ordered singular values of $\tilde{\mathbf{H}}_k$, $\mathbf{V}^{(1)} \in \mathbb{C}^{N_t \times (N_r - N_b)}$ and $\mathbf{V}^{(0)} \in \mathbb{C}^{N_t \times N_t - (N_r - N_b)}$ are the right singular matrices corresponding to the nonzero singular values and zero singular values of $\tilde{\mathbf{H}}_k$, respectively. Then select N_a out of $N_t - (N_r - N_b)$ columns of $\mathbf{V}^{(0)}$ to form $\mathbf{F}_k$. Obviously, one condition needs to be satisfied is $N_t - (N_r - N_b) \geq N_a$ or $N_t \geq N_r + N_a - N_b$. It follows that the minimum number of RF chains required at the base station is $N_t = N_r + N_a - N_b$. It should be pointed out that N_a is set to be equal N_b in [63] in order to modulate information bits based on the indices of RAs. With such an approach, each super symbol $\mathbf{x}$ at the base station is further precoded by the invertible technique, which is similar to the ZF method but it is performed at the transmitter.

5.2 Proposed MU-PQSSK System

The new multiuser downlink system investigated in this section is called multiuser precoding-aided quadrature space shift keying (MU-PQSSK), whose advantage is to reduce the minimum number of TAs required at the base station. The new system is obtained by designing matrix $\mathbf{F}$ with the following structure:

$$\mathbf{F} = [\mathbf{F}_1 \quad j\mathbf{F}_1 \quad \mathbf{F}_2 \quad j\mathbf{F}_2 \cdots \mathbf{F}_K \quad j\mathbf{F}_K] \tag{5.7}$$

Compared to (5.4), the objective in designing the new precoding matrix changes to

$$\mathbf{HF} = \begin{bmatrix} \mathbf{H}_1\mathbf{F}_1 & j\mathbf{H}_1\mathbf{F}_1 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{H}_2\mathbf{F}_1 & j\mathbf{H}_2\mathbf{F}_1 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_3\mathbf{F}_2 & j\mathbf{H}_3\mathbf{F}_2 & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_4\mathbf{F}_2 & j\mathbf{H}_4\mathbf{F}_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{5.8}$$

To satisfy (5.8), $\mathbf{F}_k$ must belong to the null space of $\tilde{\mathbf{H}}_k$, where

$$\tilde{\mathbf{H}}_k = [\mathbf{H}_1^T \cdots \mathbf{H}_{2k-2}^T \mathbf{H}_{2k+1}^T \cdots \mathbf{H}_{2K}^T]^T \in \mathbb{C}^{(N_r-2N_b)\times N_t} \tag{5.9}$$

Using SVD as in (5.6), the precoder matrix $\mathbf{F}_k$ can be obtained and the dimension condition is $N_t - (N_r - 2N_b) \geq N_a$ or $N_t \geq N_r + N_a - 2N_b$. Thus, the minimum number of RF chains at the base station is $N_t = N_r + N_a - 2N_b$, which is N_b less than that in the MU-PSSK system.

For detection, without loss of generality, consider user 1 and user 2. With the precoder matrix $\mathbf{F}$ designed to make $\mathbf{HF}$ as in (5.8), the received signals at user 1 and user 2 can be expressed as

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{F}_1(\mathbf{x}_1 + j\mathbf{x}_2) + \mathbf{n}_1 \tag{5.10}$$

$$\mathbf{y}_2 = \mathbf{H}_2\mathbf{F}_1(\mathbf{x}_1 + j\mathbf{x}_2) + \mathbf{n}_2 \tag{5.11}$$

ML Detection:

The maximum likelihood (ML) detection is as follows:

$$\hat{\mathbf{x}}_1 = \arg \min_{\forall \mathbf{x}_1, \mathbf{x}_2} \left\| \mathbf{y}_1 - \mathbf{H}_1 \mathbf{F}_1 (\mathbf{x}_1 + j\mathbf{x}_2) \right\| \tag{5.12}$$

$$\hat{\mathbf{x}}_2 = \arg \min_{\forall \mathbf{x}_1, \mathbf{x}_2} \left\| \mathbf{y}_2 - \mathbf{H}_2 \mathbf{F}_1 (\mathbf{x}_1 + j\mathbf{x}_2) \right\| \tag{5.13}$$

The complexity of the ML detection for each user is proportional to N_a^2. The next detection method is based on ZF principle and has complexity proportional to N_a.

ZF Detection:

For user 1:

1. Compute

$$\tilde{\mathbf{y}}_1 = \mathrm{Re}\{(\mathbf{H}_1 \mathbf{F}_1)^{-1} \mathbf{y}_1\} \tag{5.14}$$

2. Detect the active index:

$$\hat{m} = \arg \max_{m \in \{1,2,\cdots,N_a\}} \tilde{y}_{1,m} \tag{5.15}$$

For user 2:

1. Compute

$$\tilde{\mathbf{y}}_2 = \mathrm{Im}\{(\mathbf{H}_2 \mathbf{F}_1)^{-1} \mathbf{y}_2\} \tag{5.16}$$

2. Detect the active index:

$$\hat{n} = \arg \max_{n \in \{1,2,\cdots,N_a\}} \tilde{y}_{2,n} \tag{5.17}$$

where $\tilde{y}_{i,\ell}$ is the ℓ-th element of the column vector $\tilde{\mathbf{y}}_i$, $i = 1, 2$. Note that $\mathbf{H}_1\mathbf{F}_1$ and $\mathbf{H}_2\mathbf{F}_1$ are invertible when they are square matrices (i.e., $N_a = N_b$). In case $N_a \neq N_b$, the pseudo inverse can be used.

5.3 Performance Analysis

Since entries of $\mathbf{T}_1 = \mathbf{H}_1\mathbf{F}_1$ in (5.10) are i.i.d complex Gaussian random variables with zero mean and unit variance, the received signal for user 1 can be rewritten by

$$\mathbf{y}_1 = \sqrt{\frac{E_s}{2}}\mathbf{T}_1\mathbf{s} + \mathbf{n}_1 \tag{5.18}$$

where $\mathbf{s} = \mathbf{x}_1 + j\mathbf{x}_2$ and $\mathrm{E}\{\mathbf{s}\mathbf{s}^H\} = 2$. Given the transmit power for the link between the base station and user 1 is E_s, the signal-to-noise ratio (SNR) is defined as E_s/N_0.

ML Detection:

At user 1, the ML detection in (5.12) is rewritten by

$$\hat{\mathbf{s}} = \arg\min_{\forall \mathbf{s}} \left\| \mathbf{y}_1 - \sqrt{\frac{E_s}{2}}\mathbf{T}_1\mathbf{s} \right\| \tag{5.19}$$

Using the same method as in [48], the average symbol error probability of the proposed MU-PQSSK system can be evaluated using the following asymptotically tight union bound

$$P[\text{error}] \leq \frac{1}{2^\lambda} \sum_{m=1}^{2^\lambda} \sum_{\substack{n=1 \\ n\neq m}}^{2^\lambda} \Pr\left(\mathbf{s}_m \to \mathbf{s}_n\right), \tag{5.20}$$

where $\lambda = 2\log_2(N_a)$ and there are 2^λ possible transmitted vectors. The quantity $\Pr\left(\mathbf{s}_m \to \mathbf{s}_n\right)$ is the pairwise error probability (PEP), which is the probability of the event that the transmitter sends vector $\mathbf{s}_m$, but the receiver decides on $\mathbf{s}_n$. The pairwise error probability can be calculated as

$$\Pr(\mathbf{s}_m \to \mathbf{s}_n) = \mu_{m,n}^{N_b} \sum_{k=0}^{N_b-1} \binom{N_b - 1 + k}{k} [1 - \mu_{m,n}]^k \tag{5.21}$$

where $\mu_{m,n} = \left(1 - \sqrt{(\bar{\gamma}_{m,n}/2)/(1 + \bar{\gamma}_{m,n}/2)}\right)/2$ and $\bar{\gamma}_{m,n}$ is the mean of an exponential random variable whose value depends on the positions of the real and imaginary parts in vectors $\mathbf{s}_m, \mathbf{s}_n$ [48].

It is noted that the average symbol error probability in (5.20) is calculated with the transmitted symbol vector containing information bits for both user 1 and user 2. Given that the information bits of two users are statistically independent, the symbol error probability of each user is $P_{\text{user 1}} = P_{\text{user 2}} = \frac{1}{2}P[\text{error}]$.

ZF Detection:

At user 1, the received symbol stream is processed by $\mathbf{G} = \sqrt{\frac{2}{E_s}}\mathbf{T}_1^{-1}$. The matrix $\mathbf{T}_1$ is invertible when $N_a = N_b$. For simplicity, the following analysis only considers the case that $\mathbf{T}_1$ is the square matrix. In case $N_a \neq N_b$, the analysis can be extended straightforwardly by using pseudo inverse. The transmitted data recovered at user 1 using a ZF receiver can be expressed as

$$\mathbf{z} = \sqrt{\frac{2}{E_s}}\mathbf{T}_1^{-1}\mathbf{y}_1 = \mathbf{s} + \sqrt{\frac{2}{E_s}}\mathbf{T}_1^{-1}\mathbf{n}_1 \tag{5.22}$$

Taking the real part of $\mathbf{z}$ gives

$$\mathrm{Re}(\mathbf{z}) = \mathrm{Re}\Big\{\mathbf{s} + \sqrt{\frac{2}{E_s}}\mathbf{T}_1^{-1}\mathbf{n}_1\Big\} = \mathbf{x}_1 + \mathrm{Re}\Big\{\sqrt{\frac{2}{E_s}}\mathbf{T}_1^{-1}\mathbf{n}_1\Big\} \tag{5.23}$$

Provided that the detection process is performed using (5.14) and (5.15), a symbol error occurs when $\mathrm{Re}(\mathrm{z}_m) < \mathrm{Re}(\mathrm{z}_n)$ or equivalently $1 + \mathrm{Re}\Big\{\sqrt{\frac{2}{E_s}}[\mathbf{T}_1^{-1}]_m\mathbf{n}_1\Big\} < \mathrm{Re}\Big\{\sqrt{\frac{2}{E_s}}[\mathbf{T}_1^{-1}]_n\mathbf{n}_1\Big\}$, where z_m is the m-th element of the column vector $\mathbf{z}$ and $[\mathbf{T}_1^{-1}]_m$ is the m-th row of the square matrix $\mathbf{T}_1^{-1}$. With straightforward manipulations, it can be shown that $\mathrm{Re}\Big\{\sqrt{\frac{2}{E_s}}([\mathbf{T}_1^{-1}]_n - [\mathbf{T}_1^{-1}]_m)\mathbf{n}_1\Big\} > 1$. Given the statistical distribution of $\mathbf{n}_1$, it can be shown that $\mathrm{Re}\Big\{\sqrt{\frac{2}{E_s}}([\mathbf{T}_1^{-1}]_n - [\mathbf{T}_1^{-1}]_m)\mathbf{n}_1\Big\} \sim \mathcal{N}\left((0,\, \|[\mathbf{T}_1^{-1}]_n - [\mathbf{T}_1^{-1}]_m\|^2\frac{N_0}{E_s}\right)$. Hence, the instantaneous PEP is expressed as

$$\Pr\left(\mathbf{x}_1 \to \tilde{\mathbf{x}}_1 | \mathbf{T}_1\right) = Q(\sqrt{\gamma}) \tag{5.24}$$

where

$$\gamma = \frac{E_s/N_0}{\Big\|[\mathbf{T}_1^{-1}]_m - [\mathbf{T}_1^{-1}]_n\Big\|^2} = \frac{E_s/N_0}{\mathbf{u}^H(\mathbf{T}_1^H\mathbf{T}_1)^{-1}\mathbf{u}} \tag{5.25}$$

The column vector $\mathbf{u} \in \mathbb{R}^{N_a}$ in (5.25) comprises of zero-elements except for m-th and n-th elements which have values 1 and -1, respectively.

Since γ is a random variable, one needs to obtain the distribution of γ in order to find the average PEP. Define matrix $\mathbf{Z} = 2\mathbf{T}_1^H\mathbf{T}_1$, which is a complex Wishart matrix [68] and has the following distribution:

$$\mathcal{W}_{N_a}(N_b, \mathbf{\Sigma_z}) = \frac{\mathrm{e}^{\mathrm{Tr}(-\mathbf{\Sigma_z}^{-1}\mathbf{Z})}(\det\mathbf{Z})^{N_b - N_a}}{\Gamma(N_b)(\det\mathbf{\Sigma_z})^{N_b}} \tag{5.26}$$

where $\mathbf{\Sigma_z} = 2\mathbf{I}$ and $\Gamma(N_b) = (N_b - 1)!$ is the Gamma function.

Let $\gamma = \gamma_0 b$, where $\gamma_0 = \frac{E_s}{2N_0}$ and $b = \frac{1}{\mathbf{u}^H \mathbf{Z}^{-1} \mathbf{u}}$. Then make use of the result that if $\mathbf{Z}$ is distributed as $\mathcal{W}_{N_a}(N_b, \mathbf{\Sigma_z})$, then b is distributed as $\mathcal{W}_1(N_b - N_a + 1, \Sigma_b)$, where

$$\Sigma_b = \frac{1}{\mathbf{u}^H \mathbf{\Sigma}_\mathbf{z}^{-1} \mathbf{u}} = \frac{1}{\frac{1}{2}\mathbf{u}^H \mathbf{u}} = 1. \tag{5.27}$$

The probability density function (pdf) of b can be expressed as

$$f(b) = \frac{\mathrm{e}^{-\frac{b}{\Sigma_b}} b^{N_b - N_a}}{\Sigma_b (N_b - N_a)! (\Sigma_b)^{N_b - N_a}} \tag{5.28}$$

Performing a change of variable $\gamma = \gamma_0 b$ and using $\Sigma_b = 1$, the pdf of γ can be written directly as

$$f(\gamma) = \frac{\mathrm{e}^{-\frac{\gamma}{\gamma_0}} \gamma^{N_b - N_a}}{\gamma_0 (N_b - N_a)! (\gamma_0)^{N_b - N_a}} \tag{5.29}$$

Applying Chernoff bound $Q(\sqrt{\gamma}) \leq \frac{1}{2}\mathrm{e}^{-\frac{\gamma}{2}}$, the average PEP $\Pr(\mathbf{x}_1 \to \tilde{\mathbf{x}}_1)$ is bounded as

$$\begin{aligned} \Pr(\mathbf{x}_1 \to \tilde{\mathbf{x}}_1) &\leq \frac{1}{2}\mathrm{E}[\mathrm{e}^{-\frac{\gamma}{2}}] = \int_0^\infty \frac{1}{2}\mathrm{e}^{-\frac{\gamma}{2}} f(\gamma) d\gamma \\ &= \frac{1/2}{(1 + \frac{E_s}{4N_0})^{N_b - N_a + 1}} \end{aligned} \tag{5.30}$$

Finally, the average symbol error probability is bounded as

$$P[\text{error}] \leq (N_a - 1)\Pr(\mathbf{x}_1 \to \tilde{\mathbf{x}}_1) = \frac{(N_a - 1)/2}{(1 + \frac{E_s}{4N_0})^{N_b - N_a + 1}} \tag{5.31}$$

It can be seen from (5.31) that the diversity order of the proposed system with ZF detector is just 1 when $N_a = N_b$.

5.4 Simulation Results

In this section, simulation results in terms of the symbol error rate (SER) versus the average SNR of the proposed MU-PQSSK system are presented for different settings and detection methods. The number of users in all simulations is set to 12, while $N_a = N_b$. The number of TAs at the base station is set at minimum, i.e., $N_t = N_r + N_a - 2N_b = 11N_b$ for MU-PQSSK.

Figure 5.2 shows the performance of MU-PQSSK with $N_a = N_b = 4$ and $N_t = 44$. The spectral efficiency between the base station and one user is $\log_2(N_a) = 2$ bits/s/Hz. It can be

seen that the performance gain of ML detection method is very large over the ZF detection method. In particular, the ML detection is around 25 dB better than ZF detection at SER of 10^{-4}. The poor performance of the ZF detection is expected from the performance analysis in the previous section, which shows that the diversity order of ZF detection is only 1. The theoretical upper bounds are also plotted in Figure 5.2 and seen to be tight for both ML and ZF detection methods.

Similar performance comparison of the MU-PQSSK system is shown in Figure 5.3 when $N_a = N_b = 8$ and $N_t = 88$. The spectral efficiency between the base station and one user in this scenario is $\log_2(N_a) = 3$ bits/s/Hz. Compared to the previous case of $N_a = N_b = 4$, using a larger number of RAs with $N_a = N_b = 8$ leads to significantly better performance.

5.5 Summary

This chapter was concerned with downlink multiuser transmission between a base station and multiple users. A novel scheme, called multiuser precoding-aided quadrature space-shift keying (MU-PQSSK) was developed to mitigate the multiuser interference and reduce the minimum number of transmit antennas required at the base station. In addition to the ML detection, the low-complexity ZF detection method was also presented for the proposed system.

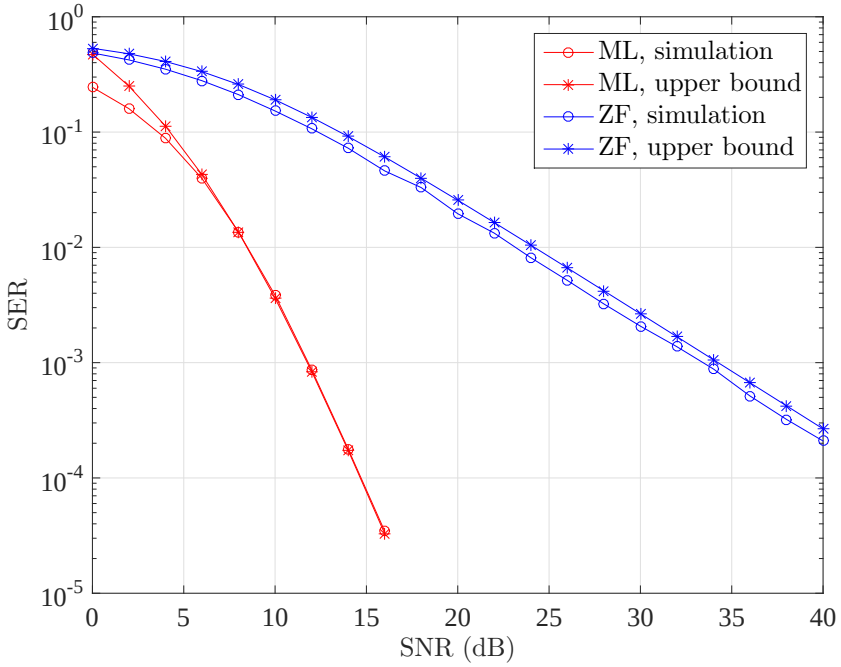

Figure 5.2 Performance of MU-PQSSK with ML and ZF detection schemes: $N_a = N_b = 4$.

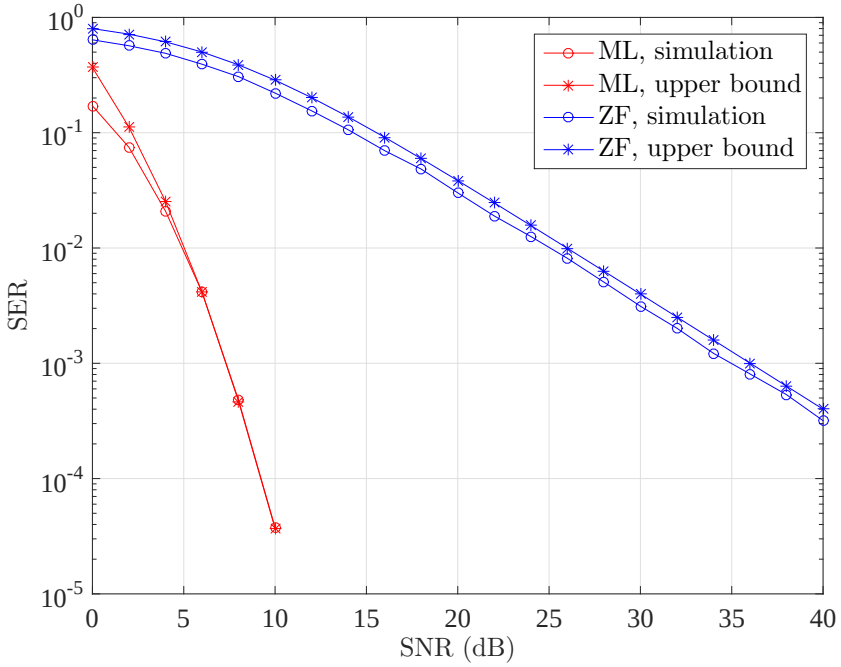

Figure 5.3 Performance of MU-PQSSK with ML and ZF detection schemes: $N_a = N_b = 8$.

6. Conclusions and Suggestions for Further Studies

6.1 Conclusions

This thesis focused mainly on developing and analyzing improved SM transmission techniques for wireless communication systems operating on frequency-flat as well as frequency-selective fading channels. The main contributions and conclusions of this thesis are summarized below.

In Chapter 3, a novel transmission scheme for a MIMO system was first developed by combining SM and a high-rate space time block code. Focusing on a system that operates with only 2 active transmit antennas, i.e., two RF chains, it was demonstrated that the proposed scheme performs better than previously-proposed schemes that are based on either Alamouti STBC or block error-control coding. A simplified ML detection of the proposed scheme was also suggested. Second, a new SSK-based modulation scheme was developed by multiplexing in-phase and quadrature GSSK streams and optimizing the signal transmitted over active antennas. Given the same number of transmit antennas, the proposed scheme doubles the transmission rate of GSSK and it achieves a higher transmission rate than BiSSK at a negligible performance loss. Third, an improved quadrature spatial modulation was proposed to enhance the spectral efficiency or error performance of a multiple-antenna communication system. This is achieved by increasing the number of spatial constellation points for each antenna dimension as well as sending an additional modulation symbol in the same symbol duration. The error performance advantage of the proposed IQSM over other existing SM schemes under the same spectral efficiency is thoroughly demonstrated by simulation and theoretical results. Lastly, the constellation design for quadrature spatial

modulation was obtained by analyzing the union upper bound on the error probability of QSM under the ML detection. Overall, the contributions in this chapter extend current knowledge in SM design for frequency-flat fading channels and offer attractive design options for making tradeoff between performance and complexity.

In Chapter 4, the SM-OFDM-GLCP scheme was developed by combining SM and OFDM in a system that employs multiple transmit antennas and operates over a frequency-selective Rayleigh fading channel. The proposed scheme implements subcarrier grouping, linear constellation precoding for each subcarrier group and selecting one transmit antenna for each subcarrier group. It was shown that the proposed SM-OFDM-GLCP achieves both diversity and coding gains which is not the case for the conventional SM-OFDM. Compared to OFDM-GLCP, the proposed scheme enhances the transmission rate by sending more information bits by means of antenna indices. Next, performance comparison between IM-based OFDM and a flexible version of OFDM, named precoded-OFDM-MConst (precoded OFDM with multiple constellations), has been performed to give a clear understanding of the IM-based OFDM system. Thanks to the multipath diversity gain achieved with precoding and enjoyed by all the information bits, the precoded-OFDM-MConst was shown to significantly outperform precoded-IM-based-OFDM systems over frequency-selective Rayleigh fading channels. The joint design of OFDM and SM presented in this chapter is therefore very attractive for realizing reliable broadband wireless communications.

Chapter 5 developed a novel downlink multiuser transmission method that is based on in-phase and quadrature SSK. The proposed method allows to reduce the minimum number of transmit antennas at the base station while keeping the complexity at the receiver low by eliminating multiuser interference. In addition to the ML detection, the low-complexity ZF detection method is also presented. The theoretical upper bounds on the performance of ML and ZF detection methods are also obtained and shown to be tight when compared with simulation results. Given the very large performance advantage of the ML detection over ZF detection and being only slightly more complicated (thanks to the downlink precoder design), it is strongly recommended that the ML detection is used with the proposed downlink multiuser transmission method.

6.2 Future Studies

The main objective of this thesis was to improve the transmission techniques for SM systems. On the path toward achieving this objective, several issues arose that should be interesting for further studies. These issues are elaborated next.

- To enhance the spectral efficiency of the conventional SM over frequency-flat fading channels, using two active antennas is very attractive because it increases the number of codewords (i.e., the number of bits conveyed by antenna index) at the expense of a slight increase in hardware complexity. A well-known transmission technique suitable for this 2-active-antenna system is STBC-SM [35]. However, the major disadvantage of STBC-SM is that it must optimize the rotation angles between codebooks. When the number of transmit antennas increases, a large number of rotation angles needs to be optimized and the coding gain of system is gets worse. Therefore, it is interesting to investigate new transmission schemes such that those rotation angles are not needed while the coding gain of system is optimized. For example, making use of the in-phase and quadrature channels together with space-time coding can increase the number of codewords without relying on the rotation angles.

- Spatial modulation uses spatial dimension (i.e., antenna index) to extend dimensions of the signal constellation. Hence, the spectral efficiency of SM is increased compared to the case of point-to-point single-antenna system. It is interesting to find out a different way to extend the signal dimensions of the conventional SM system.

- For the downlink multiuser transmission scheme developed in Chapter 5, the even-indexed and odd-indexed users are distinguished by means of in-phase and quadrature SSK. The advantage of using quadrature SSK in this scheme is to reduce the minimum number of transmit antennas required at base station with the aid of precoder matrix to mitigate multiuser interference. To enhance the data rate of this scheme, it is interesting to investigate SM instead of SSK for each user (i.e., exploiting QSM instead of quadrature SSK). Also, simplified detection methods need to be investigated to yield low decoding complexity.

References

[1] H. H. Nguyen, *Spatial Modulation: Introduction and Research Problems*. Internal Report, Department of Electrical and Computer Engieering, University of Saskatchewan, 2014.

[2] H. H. Nguyen, *Introduction to OFDM*. EE456 Lecture Notes, University of Saskatchewan, 2013.

[3] M. Maleki, H. R. Bahrami, and A. Alizadeh, "Constellation design for spatial modulation," in *Proc. IEEE Int. Conf. Commun.*, pp. 2739–2743, June 2015.

[4] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[5] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[6] M. Simon and M. S. Alouini, *Digital Communication over Fading Channels*. Wiley-IEEE Press, 2005.

[7] G. Giannakis, Z. Liu, X. Ma, and S. Zhou, *Space-Time Coding for Broadband Wireless Communications*. Wiley-Interscience, 2007.

[8] H. H. Nguyen and E. Shwedyk, *A First Course in Digital Communications*. Cambridge University Press, 2009.

[9] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *1998 URSI International Symposium on Signals, Systems, and Electronics*, pp. 295–300, Sept. 1998.

[10] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *Signal Processing Mag.*, vol. 30, pp. 40–60, Jan. 2013.

[11] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Select. Areas in Commun.*, vol. 28, pp. 1380–1408, Dec. 2010.

[12] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial modulation - a new low complexity spectral efficiency enhancing technique," in *ChinaCom*, pp. 1–5, Oct. 2006.

[13] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 2228–2241, July 2008.

[14] A. Kalis, A. Kanatas, and C. Papadias, "A novel approach to MIMO transmission using a single RF front end," *IEEE J. Select. Areas in Commun.*, vol. 26, pp. 972–980, Aug. 2008.

[15] A. Stavridis, S. Sinanovic, M. Di Renzo, H. Haas, and P. Grant, "An energy saving base station employing spatial modulation," in *17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 231–235, Sept. 2012.

[16] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Trans. on Wireless Commun.*, vol. 18, pp. 28–35, Dec. 2011.

[17] M. Di Renzo and H. Haas, "On transmit diversity for spatial modulation MIMO: Impact of spatial constellation diagram and shaping filters at the transmitter," *IEEE Trans. Veh. Technol.*, vol. 62, pp. 2507–2531, July 2013.

[18] D. Brennan, "Linear diversity combining techniques," *Proceedings of the IEEE*, vol. 91, pp. 331–356, Feb. 2003.

[19] H. Jafarkhani, *Space-time Coding: Theory and Practice*. Cambridge University Press, 2005.

[20] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, Oct. 1998.

[21] M. D. Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proceedings of the IEEE*, vol. 102, pp. 56–103, Jan. 2014.

[22] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Hoeher, "Multiple-antenna techniques for wireless communications - a comprehensive literature survey," *IEEE Communications Surveys Tutorials*, vol. 11, pp. 87–105, Second 2009.

[23] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Commun. Letters*, vol. 12, pp. 545–547, Aug. 2008.

[24] A. Mohammadi and F. M. Ghannouchi, "Single RF front-end MIMO transceivers," *IEEE Communications Magazine*, vol. 49, pp. 104–109, Dec. 2011.

[25] F. Heliot, M. A. Imran, and R. Tafazolli, "On the energy efficiency-spectral efficiency trade-off over the MIMO Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 60, pp. 1345–1356, May 2012.

[26] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys Tutorials*, vol. 13, pp. 524–540, Fourth 2011.

[27] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, pp. 1390–1436, Aug. 2011.

[28] S. Rajagopal, S. Abu-Surra, Z. Pi, and F. Khan, "Antenna array design for multi-Gbps mmwave mobile broadband communication," in *Proc. IEEE Global Telecommun. Conf.*, pp. 1–6, Dec. 2011.

[29] Q. Li and Y. P. Zhang, "Cmos T/R switch design: Towards ultra-wideband and higher frequency," *IEEE Journal of Solid-State Circuits*, vol. 42, pp. 563–570, Mar. 2007.

[30] R. B. Lai, S. F. Chao, Z. M. Tsai, J. Lee, and H. Wang, "Topology analysis and design of passive HEMT millimeter-wave multiple-port switches," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, pp. 1545–1554, July 2008.

[31] I. Y. Lee and D. Im, "Low-power SOI CMOS antenna switch driver circuit with RF leakage suppression and fast switching time," *Electronics Letters*, vol. 53, no. 5, pp. 293–294, 2017.

[32] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. on Wireless Commun.*, vol. 8, pp. 3692–3703, July 2009.

[33] M. D. Renzo, D. D. Leonardis, F. Graziosi, and H. Haas, "Space shift keying (SSK-) MIMO with practical channel estimates," *IEEE Trans. Commun.*, vol. 60, pp. 998–1012, Apr. 2012.

[34] X. Wu, H. Claussen, M. D. Renzo, and H. Haas, "Channel estimation for spatial modulation," *IEEE Trans. Commun.*, vol. 62, pp. 4362–4372, Dec. 2014.

[35] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, pp. 5536–5549, Nov. 2013.

[36] E. Basar, U. Aygolu, E. Panayirci, and H. Poor, "Space-time block coded spatial modulation," *IEEE Trans. Commun.*, vol. 59, pp. 823–832, Mar. 2011.

[37] L. Wang, Z. Chen, and X. Wang, "A space-time block coded spatial modulation from (n, k) error correcting code," *IEEE Wireless Commun. Letters*, vol. 3, pp. 54–57, Feb. 2014.

[38] B. T. Vo, H. H. Nguyen, and N. Quoc-Tuan, "High-rate space-time block coded spatial modulation," in *International Conference on Advanced Technologies for Communications (ATC)*, Oct. 2015.

[39] S. Sezginer, H. Sari, and E. Biglieri, "On high-rate full-diversity 2×2 space-time codes with low-complexity optimum detection," *IEEE Trans. Commun.*, vol. 57, pp. 1532–1541, May 2009.

[40] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Generalized space shift keying modulation for MIMO channels," in *IEEE PIMRC*, pp. 1–5, Sept. 2008.

[41] H.-W. Liang, R. Chang, W.-H. Chung, H. Zhang, and S.-Y. Kuo, "Bi-space shift keying modulation for MIMO systems," *IEEE Commun. Letters*, vol. 16, Aug. 2012.

[42] S. Fang, L. Li, S. Hu, and G. Feng, "Layered space shift keying modulation over MIMO channels," in *Proc. IEEE Int. Conf. Commun.*, pp. 1–5, June 2015.

[43] B. T. Vo and H. H. Nguyen, "Bi-generalized space shift keying over MIMO channels," *IEICE Communications Express*, Nov. 2015.

[44] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1998.

[45] M.-S. Alouini and A. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1324–1334, Sept. 1999.

[46] R. Chang, S.-J. Lin, and W.-H. Chung, "New space shift keying modulation with Hamming code-aided constellation design," *IEEE Wireless Commun. Letters*, vol. 1, pp. 2–5, Feb. 2012.

[47] A. Younis, S. Sinanovic, M. Di Renzo, R. Mesleh, and H. Haas, "Generalised sphere decoding for spatial modulation," *IEEE Trans. Commun.*, vol. 61, pp. 2805–2815, July 2013.

[48] R. Mesleh, S. Ikki, and H. Aggoune, "Quadrature spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 2738–2742, June 2015.

[49] C.-C. Cheng, H. Sari, S. Sezginer, and Y. Su, "Enhanced spatial modulation with multiple signal constellations," *IEEE Trans. Commun.*, vol. 63, pp. 2237–2248, June 2015.

[50] M. Maleki, H. R. Bahrami, A. Alizadeh, and N. H. Tran, "On the performance of spatial modulation: Optimal constellation breakdown," *IEEE Trans. Commun.*, vol. 62, pp. 144–157, Jan. 2014.

[51] M. D. Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: a survey," *IEEE Commun. Mag.*, vol. 49, pp. 182–191, Dec. 2011.

[52] P. Yang, Y. Xiao, B. Zhang, S. Li, M. El-Hajjar, and L. Hanzo, "Star-QAM signaling constellations for spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 63, pp. 3741–3749, Oct. 2014.

[53] Z. Liu, Y. Xin, and G. B. Giannakis, "Linear constellation precoding for OFDM with maximum multipath diversity and coding gains," *IEEE Trans. Commun.*, vol. 51, pp. 416–427, Mar. 2003.

[54] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," vol. 2, pp. 294–309, Mar. 2003.

[55] B. T. Vo, H. H. Nguyen, and N. Quoc-Tuan, "Spatial modulation for OFDM with linear constellation precoding," in *International Conference on Advanced Technologies for Communications (ATC)*, Oct. 2015.

[56] N. H. Tran, H. H. Nguyen, and T. Le-Ngoc, "Subcarrier grouping for OFDM with linear constellation precoding over multipath fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 3607–3613, Nov. 2007.

[57] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 50–60, 2017.

[58] X. Zhang, H. Bie, Q. Ye, C. Lei, and X. Tang, "Dual-mode index modulation aided OFDM with constellation power allocation and low-complexity detector design," *IEEE Access*, 2017.

[59] N. Serafimovski, S. Sinanovic, M. D. Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, p. 299, Sept. 2012.

[60] M. D. Renzo and H. Haas, "Bit error probability of space-shift keying MIMO over multiple-access independent fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 3694–3711, Oct. 2011.

[61] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. on Wireless Commun.*, vol. 9, pp. 3590–3600, Nov. 2010.

[62] M. Maleki, H. R. Bahrami, and A. Alizadeh, "Layered spatial modulation for multiuser communications," *IEEE Trans. on Wireless Commun.*, vol. 15, pp. 7143–7159, Oct. 2016.

[63] Y. Chen, L. Wang, Z. Zhao, M. Ma, and B. Jiao, "Secure multiuser MIMO downlink transmission via precoding-aided spatial modulation," *IEEE Commun. Letters*, vol. 20, pp. 1116–1119, June 2016.

[64] L. He, J. Wang, J. Song, and L. Hanzo, "On the multi-user multi-cell massive spatial modulation uplink: How many antennas for each user?," *IEEE Trans. on Wireless Commun.*, vol. 16, pp. 1437–1451, Mar. 2017.

[65] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. on Wireless Commun.*, vol. 14, pp. 2156–2168, Apr. 2015.

[66] A. Garcia-Rodriguez and C. Masouros, "Low-complexity compressive sensing detection for spatial modulation in large-scale multiple access channels," *IEEE Trans. Commun.*, vol. 63, pp. 2565–2579, July 2015.

[67] H. W. Liang, R. Y. Chang, W. H. Chung, H. Zhang, and S. Y. Kuo, "Bi-space shift keying modulation for MIMO systems," *IEEE Commun. Letters*, vol. 16, pp. 1161–1164, Aug. 2012.

[68] D. A. Gore, R. W. Heath, and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Letters*, vol. 6, pp. 491–493, Nov. 2002.